



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 120365

TO: Vanessa L Ford
Location: rem/3b25/3c18
Art Unit: 1645
Tuesday, April 27, 2004

Case Serial Number: 10620049

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen 1A69
Phone: 571-272-2518 *BOB*

barbara.obryen@uspto.gov

Search Notes

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 15:46:42 ; Search time 2545 Seconds

(without alignments)
8694.647 Million cell updates/sec

Title: US-10-620-049-24

Perfect score: 741

Sequence: 1 Gattacagatgacacagac.....cctcagtcacggtctctctcg 741

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pin:*

20: em_gss_vit:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pig:*

27: em_gss_vit:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	293	39.5	616	10	BF136295
2	293	39.5	941	10	BF138189
3	292.4	39.5	669	10	BF123422
4	282	38.1	462	14	CA578969

5	282	38.1	481	14	CA575938
6	281.8	38.0	955	13	BQ947728
7	281.2	37.9	827	12	BI152061
8	281.2	37.9	1459	13	BQ918407
9	280.4	37.8	774	10	BF123744
10	277.2	37.4	891	10	BF138460
11	272.4	36.8	896	12	BI149320
12	270.2	36.5	589	11	AK007163
13	269	36.3	708	12	BI103114
14	268.2	36.2	940	10	BF169023
15	267.8	36.1	417	14	CA575119
16	267.4	36.1	429	14	CA579143
17	267.4	36.1	475	14	CA578534
18	266	35.9	830	10	BF144339
19	265.8	35.9	772	10	BF581992
20	262	35.4	830	12	BI455041
21	260.8	35.2	933	13	BQ929209
22	259.2	35.0	690	12	BG968682
23	257.2	34.7	626	10	BF582283
24	257.2	34.7	725	10	BF580940
25	257	34.7	538	14	CA578968
26	256.4	34.6	724	12	BG962137
27	256	34.5	966	10	BF578083
28	255.4	34.5	979	10	BF180947
29	254.8	34.4	506	14	CA578474
30	254.8	34.4	506	14	CA578686
31	254.2	34.3	465	14	CA578116
32	254	34.3	958	13	BUS23796
33	253	34.1	735	12	BG966898
34	253	34.1	1919	10	BF180034
35	252.6	34.1	698	10	BF584024
36	251.6	34.0	1576	11	AK007918
37	250.8	33.8	794	12	BI150371
38	250.6	33.8	739	10	BE284158
39	250.6	33.8	931	13	BUS23606
40	250.2	33.8	439	10	BE136816
41	248.4	33.5	596	10	BE310080
42	247.4	33.4	863	13	BQ952037
43	247.4	33.4	916	10	BF163514
44	247	33.3	469	14	CA574907
45	246.6	33.3	473	14	CA580167

ALIGNMENTS

RESULT 1
BF136295
LOCUS 601781413F1 NCI_CGAP_Lu30 Mus musculus cdna clone IMAGE:4009443 5',
DEFINITION 616 bp mRNA linear EST 24-OCT-2000
mRNA sequence.
ACCESSION BF136295
VERSION BF136295.1 GI:10975335
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 616)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM9245 row: j column: 04
High quality sequence stop: 613.

CA576938 K0702C06-
BQ947728 AGENCOURT
BI152061 602916512
BQ918407 AGENCOURT
BF123744 601760491
BF138460 601782916
BI149320 602848859
AK007163 Mus muscu
BI103114 602889345
BF169023 601775295
CA575119 K0710F01-
CA579143 K0733D01-
CA578534 K0725B11-
BF144339 601787158
BF581992 602099448
BI455041 603173343
BQ929209 AGENCOURT
BQ968682 602836513
BF582283 602101109
BF580940 602100636
CA578968 K0731A04-
BG962137 602836902
BF578083 602094759
BF180947 601807187
CA578474 K0724C11-
CA578686 K0727C05-
CA578116 K0718G11-
BUS23796 AGENCOURT
BG966898 602834104
BF180034 601806383
BF584024 602086264
AK007918 Mus muscu
BI150371 602915205
BE284158 601099428
BUS23606 AGENCOURT
BE136816 U956501.V
BE310080 601091545
BQ952037 AGENCOURT
BF163514 601771691
CA574907 K0622H02-
CA580167 K0747G08-

```

found through the I.M.A.G.E.E. Consortium/LLNL at:
http://image.llnl.gov
Place: LRM9245 row: k column: 04
High quality sequence stop: 637.
Location/Qualifiers
1. .941
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Czech I"
/db_xref="taxon:10090"
/clone="IMAGE:4009467"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu30"
/note="Organ: lung; Vector: pCMV-SPORT6; Site1: NotI;
Site 2: SalI; transgenic model MMT-1, expression driven by
MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"

```

ORIGIN	Query Match	39.5%;	Score 293;	DB 10;	Length 941;
	Best Local Similarity	87.9%;	Pred. NO. 1.7e-58;		
	Matches 333;	Conservative 0;	Mismatches 40;	Indels 6;	Gaps 1;
QY	368	CGGCTGGTGGTGATCCGAGGTCCTCAAGATTTCTGCTACGCAATTCGCTACGCAATTCAGTCTCTGGATGA	427		
Db	48	CTGAAGGTGTCACCTCCAGGTTACGTGACCACTGACCACTCTGGACCTGAGCTGCTGAGCCCTG	107		
QY	428	GGGCTCTAGTGAAGATTTCTGCTCAAGATTTCTGCTACGCAATTCGCTACGCAATTCAGTCTCTGGATGA	487		
Db	108	GGGCTCTAGTGAAGATTTCTGCTCAAGGCTTCTGGCTACGCAATTCAGTCTCTGGATGA	167		
QY	488	ACTGGGTGAGCAGAGGCTCGACAGGCTCTGAGTGGATTGGACCGGATTTATCTCTGGAG	547		
Db	168	ACTGGGTGAAGCAGAGGCTGGAAAGGGCTTGTAGTGGATTGGACCGGATTTATCTCTGGAG	227		
QY	548	ATGGAGATTTCTAACTACATGGAATTCGAGGGCAAGGCCATACCTACAGCAGACAAAT	607		
Db	228	ATGGAGATTTCTAACTACATGGAATTCGAGGGCAAGGCCATACCTACAGCAGACAAAT	287		
QY	608	CCTCAGCAGCAGCTACATGAGCTCAGCAGCCTGACCTCTGTGCACTCTGGGCTTATT	667		
Db	288	CCTCAGCAGCAGCTACATGCAACTCAGCAGCCTGACATCTCAGGACTCTCGGGTCTACT	347		
QY	668	CTGTGTCAGAGTCGGGTTGCTACGTT-----ATGCTATGGACTACTGGGGTCAAGGAA	721		
Db	348	CTGTGTCAGAGTTAACTACTATGGCTCTACAATGCTATGGACTACTGGGGTCAAGGAA	407		
QY	722	CTCTAGTCAACGGTCTCCTC	740		
Db	408	CCTCAGTCAACGGTCTCCTC	426		

RESULT 3
BF123422
BF123422
LOCUS
DEFINITION
669 bp mRNA linear EST 24-OCT-2000
001760623F1 NCI_CGAP_Mams Mus musculus cDNA clone IMAGE:4023749 5' ,
mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 669)
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
NIH-MGC <http://mgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

```

FEATURES
source
Location/Qualifiers
1. .616
   /organism="Mus musculus"
   /mol_type="rRNA"
   /strain="Czech II"
   /db_xref="taxon:10090"
   /clone="IMAGE:4009443"
   /tissue_type="tumor, metastatic to mammary"
   /lab_host="DH10B"
   /clone_lib="NCI CGAP Lu30"
   /note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; transgenic model WNT-1, expression driven by
MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
81. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"

```

Query Match	39.5%;	Score	293;	DB	10;	Length	616;
Best Local Similarity	87.9%;	Pred. NO.	1.4e-58;				
Matches	333;	Conservative	0;	Mismatches	40;	Indels	6;
Gaps	1;						

QY	368	CGGGTGGTGGATCCGAGGTC	CAACTGCAACAGTCTGGA	ACTGAGCTGGTGAAGCCCTG	427
DB	49	CTGAAGGTGCCACTCCAGGTT	CAGCTGCAGCAGTCTGGAC	CTGAGCTGGTGAAGCCCTG	108
QY	428	GGSCCTCAGTGRAAGATTTC	TGCAAGAGATTCTGGCTACG	CATTCAATAGCTCTTGATGA	487
DB	109	GGSCCTCAGTGAAGATTTC	TGCAAGGCTTCGGCTACG	CATTCAATAGCTCTTGATGA	168
QY	488	ACTGGGTGAAGCAGAGCCCT	GGAACAGGCTTCAGTGGAT	TGACCGGATTTATCCTGGAG	547
DB	169	ACTGGGTGAAGCAGAGCCCT	GGAAGAGGCTTCAGTGGAT	TGACCGGATTTATCCTGGAG	228
QY	548	ATGGAGATTCTAATCAATGG	GAATTTGAGGGCAAGGCCAT	TACTGACAGCAGACAAAT	607
DB	229	ATGGAGATATGACTACAAT	GGGAATTTCAAGGGCAAGGCC	ACACTGACTGCAGACAAAT	288
QY	608	CCTCCAGCAGAGCCTACAT	GCAGCTCAGCAGCCTGAC	CTCTGTGGACTCTGCGGTCTATT	667
DB	289	CCTCCAGCAGAGCCTACAT	GCAGCCTGACACTCAGACG	CTGACATCTGAGGACTCTGCGGTCTACT	348
QY	668	TCGTGCAAGATCGGGTTC	TGCTAGCTT-----ATGCTAT	GGAAGCTACTGGGTTCAGGGA	721
DB	349	TCGTGCAAGAGTTAACTA	TATGGCTTTACAATGCTAT	GGAAGCTACTGGGTTCAGGGA	408
QY	722	CCTCAGTCAACCGTCTC	CTCCTC	740	
DB	409	CCTCAGTCAACCGTCTC	CGGC	427	

RESULT 2
BF138189
LOCUS
DEFINITION
BF138189 941 bp mRNA linear EST 24-OCT-2000
601781314F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4009467 5',
mRNA sequence.
ACCESSION
BF138189
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognath.; Muridae; Murinae; Mus.
1 (bases 1 to 941)
REFERENCE
NTH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: CGAPbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM9282 row: n column: 06
 High quality sequence stop: 649.
 Location/Qualifiers
 1..669

FEATURES

source
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:4023749"
 /tissue_type="tumor, gross tissue"
 /dev_stage="7 months"
 /lab_host="DH10B"
 /clone_lib="NCI_CGAP Mams"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"

ORIGIN

Query Match 39.5%; Score 292.4; DB 10; Length 669;
 Best Local Similarity 92.2%; Pred. No. 2.1e-58;
 Matches 308; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 GATATTCAGATGACAGACTACATCTCCCTGCTCTCTCTGGGACAGAGTCACC 60
 Db 53 GATATTCAGATGACAGACTACATCTCCCTGCTCTCTCTGGGACAGAGTCACC 112
 QY 61 GTCAAGTTCAGGCAAGTCAGGACATTTAGGAATTTATTTAACTGGTATCAGCAGAAACCA 120
 Db 113 ATCAAGTTCAGGCAAGTCAGGACATTTAGGAATTTATTTAACTGGTATCAGCAGAAACCA 172
 QY 121 GACGGAAGTTCAGGCAAGTCAGGACATTTAGGAATTTATTTAACTGGTATCAGCAGAAACCA 180
 Db 173 GACGGAAGTTCAGGCAAGTCAGGACATTTAGGAATTTATTTAACTGGTATCAGCAGAAACCA 232
 QY 181 AGGTTTCAGTGGCAGTGGTCTGGAACAGATTTATTTAACTGGTATCAGCAGAAACCA 240
 Db 233 AGGTTTCAGTGGCAGTGGTCTGGAACAGATTTATTTAACTGGTATCAGCAGAAACCA 292
 QY 241 GAAGATTTGGCAGTGGTCTGGAACAGATTTATTTAACTGGTATCAGCAGAAACCA 300
 Db 293 GAAGATTTGGCAGTGGTCTGGAACAGATTTATTTAACTGGTATCAGCAGAAACCA 352
 QY 301 GGACCAAGTGGCAATATAAATCGTGGAGGTG 334
 Db 353 GGACCAAGTGGCAATATAAATCGTGGAGGTG 386

RESULT 4
 LOCUS CA578969 462 bp mRNA linear EST 19-NOV-2002
 DEFINITION K0731A05-5N NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-) cDNA Library (Long) Mus musculus cDNA clone NIA.K0731A05 IMAGE:30076516 5', mRNA sequence.
 CA578969
 CA578969.1 GI:25127360
 EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (Bases 1 to 462)
 Piao,Y., Kargul,G.J., Dudekula,D.B., Qian,Y., Luo,A., Carter,M.G., Aiba,K., Taub,D., Longo,D.L., Keller,J. and Ko,M.S.H.
 Systematic Analyses of NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-) cDNA Library (Long)

JOURNAL COMMENT

Unpublished (2001)
 Other ESTs: K0731A05-3
 Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@igsun.grc.nia.nih.gov
 Plate: K0731 row: A column: 05
 Seq primer: M13 Reverse
 High quality sequence stop: 462
 POLYA=No

FEATURES

source
 Location/Qualifiers
 1..462
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6Ncr"
 /db_xref="niaEST:K0731A05-5N"
 /db_xref="taxon:10090"
 /clone="NIA:K0731A05 IMAGE:30076516"
 /tissue_type="Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)"
 /dev_stage="Age approx. 10 weeks old"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-) cDNA Library (Long)"
 /note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (<http://igsun.grc.nia.nih.gov/cDNA>). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Drs. Dennis Taub, Dan Longo (National Institute on Aging, USA), Jonathan Keller (National Cancer Institute, USA). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen]:
 5'-pGATAGTTCAGTGGCAGGCGCCCTTTTCTTTTCTTTT-3' from 0.9 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker LB-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-5. The products were purified by phenol/chloroform and NotI enzymes. The cDNAs were digested with SalI and NotI and ligated into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.1 kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN

Query Match 38.1%; Score 282; DB 14; Length 462;
 Best Local Similarity 85.1%; Pred. No. 5.1e-56;
 Matches 315; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 368 CCGGTGGTGGATCCGAGTCCAACTGCAACAGTCTGGACCTGGAGTGGAGCTG 427
 Db 76 CTGAAGGTGTCCACTCCAGGTTTCAGTGGCAGAGTCTGGACCTGGAGTGGAGCTG 135
 QY 428 GGGCCTCAGTGAAGATTTCTTGCAGAGATTTCTGGCTACGCATTCAATAGCTTTGGATGA 487
 Db 136 GGGCCTCAGTGAAGATTTCTTGCAGAGTCTTGGCTACGCATTCAATAGCTTTGGATGA 195
 QY 488 ACTGGGTGAAGCAGAGGCTGCAAGGCTCTTGGTGAATGGACGGATTTATCTCTGGAG 547
 Db 196 ACTGGGTGAAGCAGAGGCTGCAAGGCTCTTGGTGAATGGACGGATTTATCTCTGGAG 255
 QY 548 ATGGAGATTTCTTAACATCAATGGGAATTCGAGGGCAAGCCATCTACTGACAGACACAAT 607
 Db 256 ATGGAGATTTCTTAACATCAATGGGAATTCGAGGGCAAGCCATCTACTGACAGACACAAT 315
 QY 608 CTCCAGCAGACCTTACATGAGCTTCAGAGCTTCACCTCTGTGCACTCTGGGTCTATT 667

Db 316 CTTCCAGCAGCCATACATGCACTCAGCAGCCCTGACATCTGAGGACTCTGGGCTCTACT 375
 QY 668 TCTGTGCAAGATCGGGGTGCTACCTTATGCTATGAGACTCTGGGTCAGGAACCTCAG 727
 Db 376 TCTGTGCAAGAGACTACGCTAGCTATAGTGTGCTTACTGGGCGCCAGGAGCTCTGG 435
 QY 728 TCACCGTCTC 737
 Db 436 TCACGTCTC 445

RESULT 5
 CAS76938
 LOCUS
 DEFINITION
 K0702C06-5N NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)
 cDNA Library Mus musculus cDNA clone NIA:K0702C06
 IMAGE:30073757 5', mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 481)
 Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Carter, M.G.,
 Aiba, K., Taub, D., Longo, D.L., Keller, J., and Ko, M.S.H.
 Systematic Analyses of NIA Mouse Hematopoietic Stem Cell
 (Lin-/c-Kit-/Sca-1-) cDNA Library (Long)
 Unpublished (2001)
 Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Caswell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@lgsun.grc.nia.nih.gov
 Plate: K0702 row: C column: 06
 Seq primer: M13 Reverse
 High quality sequence stop: 481
 POLYA=No.

FEATURES
 source
 1. .481
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6Ncr"
 /db_xref="taxon:10090"
 /clone="NIA:K0702C06 IMAGE:30073757"
 /tissue_type="Hematopoietic Stem Cell"
 (Lin-/c-Kit-/Sca-1-)
 /dev_stage="Age approx. 10 weeks old"
 /lab_host="PH108"
 /clone_lib="NIA Mouse Hematopoietic Stem Cell"
 (Lin-/c-Kit-/Sca-1-) cDNA Library (Long)
 /note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
 NotI; Mouse cDNA project by the Laboratory of Genetics,
 National Institute on Aging (NIA), Intramural Research
 Program, NIH (http://lgsun-grc.nia.nih.gov/cDNA). This is
 a long-transcript enriched cDNA library (Ref. Genome Res.
 11: 1553-1558 (2001) [PMID: 11544199]). Total RNAs were
 obtained from Drs. Dennis Taub, Dan Longo (National Cancer
 Institute on Aging, USA), Jonathan Keller (National Cancer
 Institute, USA). Double-stranded cDNAs were synthesized
 with an oligo(dT) primer [Invitrogen]:
 5'-pGACTAGTCTAGATCGGCGCGCCCTTTT-3' from
 0.9 ug of total RNA, treated with T4 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to Loner-linker LL-Sal4, purified by phenol/chloroform, and
 separated from free linkers by Centricon 100. Then, the
 cDNAs were amplified by long-range high fidelity PCR using
 Ex Taq polymerase (Takara) with a primer Sal4-S. The
 products were purified by phenol/chloroform and Centricon
 100. The cDNAs were digested with SalI and NotI enzymes
 and cloned into SalI/NotI site of pSPORT1 plasmid vector.

The DH10B E. coli host was transformed with the ligation
 mixture by the standard chemical method. The average
 insert size is about 2.1 kb. The library was constructed
 by Yulan Piao (NIA)."

ORIGIN
 Query Match 38.1%; Score 282; DB 14; Length 481;
 Best Local Similarity 85.1%; Pred. No. 5.2e-56;
 Matches 315; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
 QY 368 CGGTGGTGGATCCGAGGTCCAACTGCAAGTCTGACCTGAGCTGGTGAACCTG 427
 Db 76 CTGAAGGTGTCCACTCCCAAGGTTCAAGCTGCAGCAGTCTGACCTGAGCTG 135
 QY 428 GGGCTCAGTCAAGATTCTCTGCAAGATTCTGCTACGCAATCAATAGCTCTTGTGATGA 487
 Db 136 GGGCTCAGTCAAGATTCTCTGCAAGCTTCTGCTACGCAATCAATAGCTCTTGTGATGA 195
 QY 488 ACTGGTGAAGCAGAGCGCTGGAACAGGGTCTTGAGTGGATTGACGGAATTCCTGGAG 547
 Db 196 ACTGGTGAAGCAGAGCGCTGGAAGAGGGTCTTGAGTGGATTGACGGAATTCCTGGAG 255
 QY 548 ATGGAGATTTCAACTACAAATGGGAAATTCGAGGCAAGGCCATCTACAGCAGACAAAT 607
 Db 256 ATGGAGATTTCAACTACAAATGGGAAATTCGAGGCAAGGCCATCTACAGCAGACAAAT 315
 QY 608 CTTCCAGCAGCCATACATGCAAGTCTGAGCAGCTGAGCTCTGAGGACTCTGCGGTCTATT 667
 Db 316 CTTCCAGCAGCCATACATGCAAGTCTGAGCAGCTGAGCTCTGAGGACTCTGCGGTCTATT 375
 QY 668 TCTGTGCAAGATCGGGTCTGAGTATGCTATGAGTCTGAGGCTCAAGGAACCTCAG 727
 Db 376 TCTGTGCAAGATCGGGTCTGAGTATGCTATGAGTCTGAGGCTCAAGGAACCTCAG 435
 QY 728 TCACGTCTC 737
 Db 436 TCACGTCTC 445

RESULT 6
 BQ947728
 LOCUS
 DEFINITION
 BQ947728 955 bp mRNA linear EST 21-AUG-2002
 IMAGE:6394985 5', mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 955)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM1388 row: 0 column: 18
 High quality sequence stop: 712.
 Location/Qualifiers
 1. 955
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:6394985"


```

/clone="IMAGE:4023550"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH108"
/clone_lib="NCI CGAP Mams"
/note="organ: mammary; Vector: pcwv-sport6; Site 1: Salt;
site 2: Not; Cloned unidirectionally. Primer: Oligo dr.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NTH"

```

ORIGIN

Query Match	37.8%;	Score 280.4;	DB 10;	Length 774;
Best Local Similarity	91.9%;	Pred. No. 1.5e-55;		
Matches 307;	Conservative	0;	Mismatches 26;	Indels 1;
Gaps 1;				
Qy	1	GATATT	CAGATGACACAGACTACATCCTCCCTGTCCTCTCTGGAGACAGAGTCACC	60
Db	50	GATATCCAGATGACACAGACTACATCCTCCCTGTCCTCTCTGGAGACAGAGTCACC	109	
Qy	61	GTCAAGTTGCAGGCGAAGTCAGGACATTAGGAATTATTAACTGGTATCAGCAGAAACCA	120	
Db	110	ATCAGTTGCAGGCGAAGTCAGGACATTAGCAATTTTAACTGGTATCAGCAGAAACCA	169	
Qy	121	GACGGAACCTGTTAAATTCCTGATCTACTACACATCAAGATTACTGCCAGGAGTCCCCATCA	180	
Db	170	GACGGAACCTGTTAACTCCTGATCTATTACACATCAACATTACAAACAGGAGTCCCCATCA	229	
Qy	181	AGGTTCAAGTGGCGAGTGGGTCTGGAAACAGATTATTCCTCTCACCATTACAACTTGAGCAG	240	
Db	230	AGGTTCAAGTGGCGAGTGGGTCTGGAAACAGATTATTCCTCTCACCATTAGCAACCTTGAGCAA	289	
Qy	241	GAGAGATTTCGCACTTACTTTTTCGCAACGGGCAATACGGCTCCGTCGACGTTCCGTTGGGA	300	
Db	290	GAGAGATA-TGCCACTTACTTTTTCGCCACAGGGTTAAGCGGTTCTCTCGACGTTCCGTTGGGA	348	
Qy	301	GGCACCAAGCTGAAAATAAAAAACGTGGTGGAGGTG	334	
Db	349	GGCACCAAGGTGGAAATCAAAACGGGCTGATGCTG	382	

[illegible]

CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LRAM9249 row: o column: 04
High quality sequence rows: 581.

FEATURES SOURCE

```

Location/Qualifiers
1. .891
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Czech II"

```

```

/db_xref="taxon:10090"
/clone="IMAGE:4011099"
/tissue_type="tumor, metastatic to mammary"
/lab_host="Czech II"
/clone_lib="NCI_CGAP_Lu30"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; transgenic model WNT-1, expression driven by
MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"

ORIGIN
Query Match      37.4%; Score 277.2; DB 10; Length 891;
Best Local Similarity 88.8%; Pred. No. 9.2e-55;
Matches 324; Conservative 0; Mismatches 34; Indels 7; Gaps 2;

QY 383 CCGAGGTCAACCTGCAACAGCTCGGACCTGGAGCTGGTGAAGCTCGGGCTCAGTGAAGA 442
DB 61 CCAGGTTTCAGCTGCAGCAGCTCGGACCTGAGCTGGTGAAGCTCGGGCTCAGTGAAGA 120
QY 443 TTTCTGCGAAAGATTCGTGCTAGCAGCTCAATAGCTCTTGGATGAAGTGGGTGAAGCAGA 502
DB 121 TTTCTGCGAAGGCTTCTGGCTAGCAGCTCAATAGCTCTTGGATGAAGTGGGTGAAGCAGA 180
QY 503 GGCTCGACAGGGTCTT-GAGTGGATTGGACGAGTTTATCTGGAGATGGAGATTCTAAC 561
DB 181 GGCTCGAAGGCTTCTGGATGGATTGGACGAGTTTATCTGGAGATGGAGATTCTAAC 240
QY 562 TACAATGGGAATTCGAGGCAAGGCCATCTGACAGCAGACAAATCTCCAGCAGCAGCC 621
DB 241 TACAATGGGAATTCGAGGCAAGGCCATCTGACAGCAGACAAATCTCCAGCAGCAGCC 300
QY 622 TACATCAGCTCAGCAGCTGAGCTCTGAGGCTCTGCGGTCTATTCTGTCGACAGATCG 681
DB 301 TACATCAGCTCAGCAGCTGAGCTCTGAGGCTCTGCGGTCTATTCTGTCGACAGATCG 360
QY 682 GGGTTGCTAGCTT-----ATGCTATGAGTACTTGGGGTCAAGGAACCTCAGTCCCGTC 735
DB 361 AACTACTATGGCTTCTACATGCTATGAGTACTTGGGGTCAAGGAACCTCAGTCCCGTC 420
QY 736 TCCTC 740
DB 421 TCCGC 425

RESULT 11
LOCUS B1149320
DEFINITION B1149320 896 bp mRNA linear EST 05-JUL-2001
ACCESSION B1149320.1 GI:14609321
VERSION B1149320
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11063 row: e column: 16
High quality sequence stop: 862.
Location/Qualifiers

FEATURES
1. 896
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Czech II"
/db_xref="taxon:10090"
/clone="IMAGE:5012343"
/tissue_type="spontaneous tumor, metastatic to mammary.
Stem cell origin."
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu29"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

ORIGIN
Query Match      36.8%; Score 272.4; DB 12; Length 896;
Best Local Similarity 84.8%; Pred. No. 1.3e-53;
Matches 319; Conservative 0; Mismatches 51; Indels 6; Gaps 1;

QY 368 CCGTGTGTGTGGATCCGAGGTCCAACTGCAACAGCTCGGACCTGAGCTGGTGAAGCCTG 427
DB 81 CTGAAGGTGTCCAACTCCAGGTTTCAGCTGCAGCAGCTCTGGACCTGAGCTGTGAAGCCTG 140
QY 428 GGGCTCAGTGAAGATTTCTGCAAGATTTCTGGTACGCAATCAATAGCTCTTGGATGA 487
DB 141 GGGCTCAGTGAAGATTTCTGCAAGATTTCTGGTACGCAATCAATAGCTCTTGGATGA 200
QY 488 ACTGGGTGAAGCAGAGCCCTGGACAGGGTCTTGGTGGATTGGACGGATTATCTCTGGAG 547
DB 201 ACTGGGTGAAGCAGAGCCCTGGAAAGGGTCTTGGTGGATTGGACAGATTATCTCTGGAG 260
QY 548 ATGAGATTTCTTAATCAATGGGAATTCGAGGGCAAGGCCATCTGACAGCAGACAAAT 607
DB 261 ATGAGATTTCTTAATCAATGGGAATTCGAGGGCAAGGCCATCTGACAGCAGACAAAT 320
QY 608 CTTCCAGCAGACCTTACATGCTCAGCTCAGCAGCCCTGACCTCTGTGGACTCTCCGCTCTATT 667
DB 321 CTTCCAGCAGACCTTACATGCTCAGCTCAGCAGCCCTAATCTGAGGACTCTCGGCTCTATT 380
QY 568 TCTGTGCAAGATCGGGTGTGCTAGCTTA-----TGCATGAGTACTCTGGGTCAAGGAA 721
DB 381 TCTGTGCAAGATCCCTGTATGTTGGTTTACTACGTTTGTGTTTCTACTGGGGCCAGGGA 440
QY 722 CTTCCAGTCAAGCTCTC 737
DB 441 CTTCCAGTCAAGCTCTC 456

RESULT 12
LOCUS AK007163
DEFINITION AK007163 689 bp mRNA linear HTC 20-SEP-2003
ACCESSION AK007163.1 GI:12840553
VERSION AK007163
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Carninci, P. and Hayashizaki, Y.
JOURNAL High-efficiency full-length cDNA cloning
MEDLINE Meth. Enzymol. 303, 19-44 (1999)
PUBMED 99279253
REFERENCE 10349636
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
TITLE Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalisation and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes

```


JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.

TITLE
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5

5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 689)

JOURNAL
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, C., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

TITLE
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

COMMENT
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGAGCGCGCGCTAATTCGAGTTAATTAATTAATTAATTCGCCGCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOUR.

FEATURES
source
1..689
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:1700110L11"
/db_xref="MGI:1910224"
/db_xref="taxon:10090"
/clone="1700110L11"
/sex="male"
/tissue_type="testis"

CDS
/clone lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
155..490
/note="unamed protein product: immunoglobulin heavy chain 6 (heavy chain of IgM) (MGDI:56448, GB|AK008342, evidence: BLASTN, 100%, match=137)
putative"
/codon_start=1
/protein_id="BAB24877.1"
/db_xref="GI:12840554"
/translation="MTGVHSGVOLQQSGPBLVKPGASVKISCKASGVAFASSSMNWVK ORPKGKLEWIGRIYPGGDINNGKFKGKATLADKSSSTAYMQLSLTSEDSAVYFC ARHSVTTTS"

ORIGIN
Query Match 36.5%; Score 270.2; DB 11; Length 689;
Best Local Similarity 92.5%; Pred. No. 3.7e-53;
Matches 284; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 372 TGTGTGTGGATCCGAGGTCCTCAACTGCAACATCTGGACCTGAGCTGGTGAAGCTGGGGC 431
DB 160 TGTGTTCACCTCCAGGTTTCAGCTGCAGCATCTGGACCTGAGCTGGTGAAGCTGGGGC 219
QY 432 CTCAGTGAAGATTTCTGCTCAAGATTTCTGGCTAGCATTCAATAGCTCTTGGATCAACTG 491
DB 220 CTCAGTGAAGATTTCTGCTCAAGATTTCTGGCTAGCATTCAATAGCTCTTGGATCAACTG 279
QY 492 GGTGAAGCAGAGGCTCGACAGGGTCTTGGATGAGTGGAGGATTTATCTGGAGATGG 551
DB 280 GGTGAAGCAGAGGCTCGACAGGGTCTTGGATGAGTGGAGGATTTATCTGGAGATGG 339
QY 552 AGATTCTTACTACATATGGGAATTCGAGGCAAGGCCATCTGACAGCAGACAAATCTCTC 611
DB 340 AGATTCTTACTACATATGGGAATTCGAGGCAAGGCCATCTGACAGCAGACAAATCTCTC 399
QY 612 CAGCAGAGCTACATGAGCTGACAGCTGACCTTCTGTGACTCTGGGCTCTATTCTG 671
DB 400 CAGCAGAGCTACATGAGCTGACAGCTGACCTTCTGTGACTCTGGGCTCTATTCTG 459
QY 672 TGCAGA 678
DB 460 TGCAGA 466

RESULT 13
BI031114
LOCUS BI031114 708 bp mRNA linear EST 26-JUN-2001
602889345F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:5044690
5', mRNA sequence.
DEFINITION
ACCESSION
VERSION BI031114.1 GI:14554007
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 708)
NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLML1122 row: i column: 11
High quality sequence stop: 706.

FEATURES
source
1..708
Location/Qualifiers

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/db_xref="FVB/N"
/clone="IMAGE:5044690"
/lab_host="NCI CGAP Kid14"
/notes="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library. |"

```

ORIGIN

```

Query Match      36.3%; Score 269; DB 12; Length 708;
Best Local Similarity 88.0%; Pred. No. 7.2e-53;
Matches 293; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 2 ATATTGAGTACACAGACTACATCTCTCCCTGTCCTCTCTGGGAGACAGTCCACG 61
Db 73 ATGTCCAGATGACACAGACTACATCTCTCCCTGTCCTCTCTGGGAGACAGTCCACCA 132
QY 62 TCAGTTCCAGGCAAGTCAGACATAGGAAATATTAACTGGTATCAGCAGAAACCCAG 121
Db 133 TCATTGCAAGTCAATGCAAAATTCGCCACTTATTGAACTGGTTTCAGCAGAGCCAG 192
QY 122 ACCGAATGTTAAATCTCTGATCTACTACATCAAGATTAATTCCTCCAGGAGTCCCAATCAA 181
Db 193 ATGGAATGTTAAATCTCTGATCTACTACATCAAGATTAATTCCTCCAGGAGTCCCAATCAA 252
QY 182 GGTTCAGTGCAGTGGTCTGGACAGATTAATTCCTCCAGGAGTCCCAATCAA 241
Db 253 GGTTCAGTGCAGTGGTCTGGACAGATTAATTCCTCCAGGAGTCCCAATCAA 312
QY 242 AAGTATTGGCACTTACTTTTGGCAACAGGCAATAGCCCTCCGTTGGAGCTTCGGTGGAG 301
Db 313 AGGATATTGGCGCTTATTTTGGCAACAGGCAATAGCCCTCCGTTGGAGCTTCGGTGGAG 372
QY 302 GCACCACTGGAATTAATTAACGTCGTGAGTG 334
Db 373 GCACCTACCTGGAATTAATTAACGTCGTGAGTG 405

```

```

RESULT 14
BF169023
LOCUS      BF169023
DEFINITION BF169023.1 GI:11049375
            Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE   1 (bases 1 to 940)
AUTHORS    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE      NIH-MGC http://www.ncbi.nlm.nih.gov/
COMMENT    National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs@emall.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LIA9265 row: m column: 23
            High quality sequence stop: 581.
            Location/Qualifiers
                1..940
                    /organism="Mus musculus"
                    /mol_type="mRNA"

```

```

/strain="Czech II"
/db_xref="taxon:10090"
/clone="IMAGE:4017214"
/tissue_type="spontaneous tumor, metastatic to mammary.
Stem cell origin."
/lab_host="DH10B"
/clone_lib="NCI CGAP Lu29"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

```

ORIGIN

```

Query Match      36.2%; Score 268.2; DB 10; Length 940;
Best Local Similarity 86.1%; Pred. No. 1.3e-52;
Matches 321; Conservative 0; Mismatches 48; Indels 4; Gaps 2;

QY 368 CGGTGGTGGTGGTCCAGGTCCAACTGCAACAGTCTGGACCTGAGCTGGTGAAGCTG 427
Db 52 CTGAAGGTGTCCAAATCCAGGTTCAGCTGCAGCAGTCTGGACCTGAGCTGGTGAAGCTG 111
QY 428 GGGCTCAGTGAAGATTTCTGCAAGATTTCTGCTAGCATTCAATAGCTCTTTGGATGA 487
Db 112 GGGCTCAGTGAAGATTTCTGCAAGATTTCTGCTAGCATTCAATAGCTCTTTGGATGA 171
QY 488 ACTGGGTGAAGCAGAGGCTCGACAGGCTCTTGAGTGAATGAGCGGATTTATCTGGAG 547
Db 172 ACTGGGTGAAGCAGAGGCTCGACAGGCTCTTGAGTGAATGAGCGGATTTATCTGGAG 230
QY 548 ATGGAGATTTCAATACATGGAATTCGAGGCGCAAGCCATCTACTGACACAGCAAAAT 607
Db 231 ATGGAGATTTCAATACATGGAATTCGAGGCGCAAGCCATCTACTGACACAGCAAAAT 290
QY 608 CTTCCAGCAGACCTTACATGCACTCAGCAGCTTGTGAGCTTGTGGAGCTTCGGTCTATT 667
Db 291 CTTCCAGCAGACCTTACATGCACTCAGCAGCTTGTGAGCTTGTGGAGCTTCGGTCTATT 350
QY 668 TCTGTGCAAGATCGGGTGTGCTAGTTATGCTATGCTAGTCTGGTCAAGGACCTCAG 727
Db 351 TCTGTGCAAGATCGGGTGTGCTAGTTATGCTATGCTAGTCTGGTCAAGGACCTCAG 407
QY 728 TCACCGTCTCCTC 740
Db 408 TCACAGTCTCCTC 420

```

```

RESULT 15
CA577519
LOCUS      CA577519
DEFINITION CA577519.1 GI:25125910
            Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE   1 (bases 1 to 417)
AUTHORS    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE      Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Carter, M.G.,
            Aiba, K., Taub, D., Longo, D.L., Keller, J. and Ko, M.S.H.
            Systematic Analyses of NIA Mouse Hematopoietic Stem Cell
            (Lin-/c-Kit-/Sca-1-) cDNA Library (Long)
            Unpublished (2001)
            Other ESTs: K0710F01-3
            Contact: Dawood B. Dudekula
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@igsun.grc.nia.nih.gov
            Plate: K0710 row: F column: 01
            Seq primer: M13 Reverse

```

High quality sequence stop: 417

FEATURES
POLYA=No. Location/Qualifiers
1. .417
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6NCr"
/db_xref="niaEST:K0710F01-5N"
/db_xref="taxon:10090"
/clone="NIA:K0710F01 IMAGE:30074556"
/tissue_type="Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-)"
/dev_stage="Age approx.10 weeks old"
/lab_hosts="DH10B"
/clone_lib="NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-) cDNA Library (Long)"
/note="vector: pSPOR1 (Invitrogen); Site_1: SalI; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (<http://lgsun.grc.nia.nih.gov/cDNA>). This is
a long-transcript-enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
obtained from Dr. Dennis Taub, Dan Longo (National Cancer
Institute on Aging, USA), Jonathan Keller (National Cancer
Institute, USA). Double-stranded cDNAs were synthesized
with an Oligo(dT) primer [Invitrogen:
5'-pGACTAGTCTAGATCGGCGCCCTTTT-3'] from
0.9 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker Lu-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pSPOR1 plasmid vector.
The DH10B E. coli host was transformed with the ligation
mixture by the standard chemical method. The average
insert size is about 2.1 kb. The library was constructed
by Yulan Piao (NIA)."

ORIGIN

Query Match 36.1%; Score 267.8; DB 14; Length 417;
Best Local Similarity 91.3%; Pred. No. 1.1e-52;
Matches 284; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy	368	CCGCTGGTGGTGGATCCAGAGTCCAACTGCAACAGTCTGGACCTGAGCTGGTGAAGCTG	427
Db	76	CTGAAGGTGTCCACTCCAGCTTCAGCTGCAGCAGTCTGGACCTGAGCTGGTGAAGCTG	135
Qy	428	GGGCTCAGTGAAGATTTCTGCAAGATTCCTGCTAGCTGCAATAGCTCTTGGATGA	487
Db	136	GGGCTCAGTGAAGATTTCTGCAAGGCTTCTGGCTAGCTGCAATAGCTCTTGGATGA	195
Qy	488	ACTGGGTGAAGCAGAGGCTGCAAGGCTTTGAGTGGATTGGACGGATTTATCCTGGAG	547
Db	196	ACTGGGTGAAGCAGAGGCTGCAAGGCTTTGAGTGGATTGGACGGATTTATCCTGGAG	255
Qy	548	ATGGAGATCTTAATCAATGGGAATTCGAGGCGAAGCCATCTACACAGACAAAT	607
Db	256	ATGGAGATCTTAATCAATGGGAATTCGAGGCGAAGCCATCTACACAGACAAAT	315
Qy	608	CCTCCAGACACACCTACATGAGCTCAGCAGCCTGACCTCTGTGACTCTCGGCTCTATT	667
Db	316	CCTCCAGACACACCTACATGAGCTCAGCAGCCTGACATCTGAGGACTCTCGGCTCTACT	375
Qy	668	TCTGTGCAAGA	678
Db	376	TCTGTGCAAGA	386

Search completed: April 26, 2004, 17:37:51
Job time : 2550 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 15:43:02 ; Search time 393 seconds
(without alignments)

8009.961 Million cell updates/sec

Title: US-10-620-049-24

Perfect score: 741
Sequence: 1 gatattcagatgacacagac.....cctcagtcacgctctctcgcg 741

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 337863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 29Jan04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002s:*

7: geneseqn2003as:*

8: geneseqn2003bs:*

9: geneseqn2003cs:*

10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	499.4	67.4	1803	3 AAAS8773	Aaas8773 DNA encod
2	499.4	67.4	1943	6 ABK13467	Abk13467 DNA encod
3	499.4	67.4	1944	7 ABX14413	Abx14413 DNA encod
4	482.8	65.2	2952	6 ABK13464	Abk13464 DNA encod
5	479.2	64.7	913	2 AAQ81076	Aaq81076 Bispecifi
6	463.8	62.6	1527	9 ADD25786	Add25786 Binding d
7	463.8	62.6	1687	9 ADE86042	Ade86042 DNA encod
8	431.8	58.3	2691	6 ABK13612	Abk13612 DNA encod
9	421	56.8	1953	8 ACF05482	Acf05482 Anti-CD7
10	421	56.8	1956	8 ACF05481	Acf05481 Anti-CD7
11	418.6	56.5	1457	3 AA39167	Aa39167 Apoptobod
12	406	54.8	1545	6 AAD32522	Aad32522 Murine CC
13	404.2	54.5	1698	3 AAZ43431	Aaz43431 Fv-antibo
14	395.8	53.4	1817	7 ACC79606	Acc79606 Plasmid p
15	394.2	53.2	1611	3 AAZ30332	Aaz30332 Nucleotid
16	391.6	52.8	795	6 ABK89855	Abk89855 Mouse sin
17	391.6	52.8	795	9 ADD25448	Add25448 Binding d
18	385.6	52.0	1637	3 AAZ89358	Aaz89358 Bispecifi
19	384.8	51.9	1065	2 AAT94963	Aat94963 R. pipien
20	384.8	51.9	1065	2 AAT94967	Aat94967 R. pipien
21	384.8	51.9	1074	2 AAT94965	Aat94965 R. pipien
22	384.8	51.9	1074	2 AAT94968	Aat94968 R. pipien
23	384.8	51.9	1086	2 AAT94966	Aat94966 R. pipien

24 384.8 51.9 1137 2 AAT94964
25 383.8 51.8 825 6 ABK89854
26 383.8 51.8 825 9 AAD25447
27 381.8 51.5 819 2 AAQ20381
28 379.2 51.2 752 4 AAF81912
29 379.2 51.2 782 4 AAF81913
30 373.2 50.4 1817 7 ACC79607
31 371.2 50.1 1612 4 AAC86562
32 371.2 50.1 1614 8 AAT43497
33 368 49.7 786 8 AAD56816
34 368 49.7 993 8 AAD56805
35 368 49.7 1017 8 AAD56800
36 367.8 49.6 848 2 AAQ41069
37 363.8 49.1 2871 3 AAZ50586
38 361.8 48.8 1653 3 AAZ43432
39 358.6 48.4 925 3 AAZ44205
40 358.6 48.4 925 3 AAZ58663
41 356.4 48.1 1299 2 AAV63620
42 356.4 48.1 1299 2 AAZ09862
43 356.4 48.1 1320 2 AAV63617
44 356.4 48.1 1320 2 AAZ09859
45 356.4 48.1 6727 2 AAV63616

ALIGNMENTS

RESULT 1
AAAS8773

ID AAAS8773 standard; DNA; 1803 BP.

XX

AC AAAS8773;

XX

DT 20-OCT-2000 (first entry)

XX

DE DNA encoding an immunotoxin comprising a CD3 binding domain and PE28.

XX

KW Immunotoxin; CD3-binding domain; Pseudomonas exotoxin A; immune system;

KW scFv(UCHT-1)-PE28; T-cell mediated disease; transplant rejection;

KW host versus graft disease; graft versus host disease;

KW bone marrow transplant; ss.

XX

OS Synthetic.

OS Mus sp.

OS Pseudomonas aeruginosa.

XX

PN WO200041474-A2.

XX

PD 20-JUL-2000.

XX

PF 13-JAN-2000; 2000WO-EP000245.

XX

PR 15-JAN-1999; 99US-00232445.

PR 25-JAN-1999; 99US-00236968.

PR 07-OCT-1999; 99US-00414134.

XX

PA (NOVS) NOVARTIS AG.

PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.

XX

PI Digan ME, Lake P, Wright RM;

XX

DR WPI; 2000-482739/42.

DR P-PSDB; AAB07461.

XX

PT Recombinant immunotoxin used for prophylaxis and treatment of T-cell

PT mediated diseases e.g. transplantation rejection.

XX

PS Disclosure; Page 55-59; 75pp; English.

XX

CC The present sequence encodes a recombinant immunotoxin comprising a CD3-binding domain and a Pseudomonas exotoxin A component. It is designated

CC scFv(UCHT-1)-PE28. The immunotoxins of the invention are used for the

CC prophylaxis or treatment of T-cell mediated diseases or conditions of the

CC immune system. They are also used to condition a patient to be
 CC transplanted with cells, or a tissue or an organ of a donor and for the
 CC prophylaxis and/or treatment of acute or chronic transplantation
 CC rejection, host versus graft disease and/or graft versus host disease in
 CC a patient to undergo a bone marrow transplant, where the CD3-bearing cell
 CC population in the patient is depleted and an inoculum comprising isolated
 CC bone marrow and/or stem-cell enriched peripheral blood cells of the donor
 CC treated with immunotoxin is inoculated into the patient
 XX
 SQ Sequence 1803 BP; 363 A; 575 C; 567 G; 298 T; 0 U; 0 Other;
 Query Match 67.48; Score 499.4; DB 3; Length 1803;
 Best Local Similarity 81.44; Pred. No. 1e-120;
 Matches 610; Conservative 0; Mismatches 121; Indels 18; Gaps 2;
 QY 1 GATATTGAGATGACACAGATACATCTCCCTCTCTGCTCTCTGGGAGACAGATCACC 60
 Db 7 GACATCCAGATGACCCAGACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 66
 QY 61 GTGAGTTGACGGGCAAGTCAGGACATAGGATATTTAACTGGTATCAGAGAAACCA 120
 Db 67 ATCAGTTGACGGGCAAGTCAGGACATAGGATATTTAACTGGTATCAGAGAAACCA 126
 QY 121 GACGGAACTGTAAATTCCTGATCTACTACATCATCAAGATTACTGCCAGGAGTCCCATCA 180
 Db 127 GATGGAACCTGTAACTCTGATCTACTACATCATCAAGATTACTGCCAGGAGTCCCATCA 186
 QY 181 AGTTTCAGTGCAGTGGTCTGGAACAGATATTCCTCTACCATTAACACCTGGAGCAG 240
 Db 187 AAGTTTCAGTGCAGTGGTCTGGAACAGATATTCCTCTACCATTAACACCTGGAGCAG 246
 QY 241 GAAGATATTGGCACTTACTTTTGGCAACAGGCAATACGCTCCGTCGACGTTCCGGTGA 300
 Db 247 GAGGATATTGCACTTACTTTTGGCAACAGGCAATACGCTCCGTCGACGTTCCGGTGA 306
 QY 301 GGCACCAAGCTGGAAATAAAGCTGGTGGAGTGGTCTGATGTTGGTGTCTCTGGCGGC 360
 Db 307 GGCACCAAGCTGGAAATAAAGCTGGTGGAGTGGTCTGATGTTGGTGTCTCTGGCGGC 366
 QY 361 GGCAGCTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420
 Db 367 AGC-----GGTGGCGGATCTGAGTGCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCT 417
 QY 421 AAGCTGGGGCTCAGTGAAGATTTCCTGCAAGATTCTGGCTACGATTCATTAATAGTCT 480
 Db 418 AAGCTGGAGCTTCAATGAAGATATCTGCAAGGCTTCTGGTTACTCACTCACTGGCTAC 477
 QY 481 TGGATGACTGGTGCAGCAGGCTGGACAGGCTCTGAGTGGATTGGACGATTTAT 540
 Db 478 ACCATGAACCTGGTGAACAGAGTATGGAAGAACCTTGGTGGATGGAGCTTATTAAT 537
 QY 541 CCTGGAGATGGAGATTCTTAACATAATGGGAAATTCGAGGCAAGGCCATCTGACAGCA 600
 Db 538 CCTTACAAAGTGTGTAGTACCTACACCAAGAGTTCAAGGCAAGGCCACATTAACGTGA 597
 QY 601 GACAAATCTTCAGACAGCTTACATGACGCTCAGCAGCTCAGTCTGTGGACTCTGG 660
 Db 598 GACAAATCTTCAGACAGCTTACATGACGCTCAGCAGCTCAGTCTGTGGACTCTGG 657
 QY 661 GTCTATTCTGTGCAAGATCGGGTGTGCTAGCTTATGCT-----ATGCACTATGG 711
 Db 658 GTCTATTCTGTGCAAGATCGGGTGTGCTAGCTTATGCTAGCTGCTAGCTGCTAGCTGCTAG 717
 QY 712 GGTCAAGAACCTCAGTCAAGCTCTCTCTC 740
 Db 718 GGCGCAGGAGCACAGGTCACCGTCTCTCTC 746
 RESULT 2
 ABK13467
 ID ABK13467 standard; DNA; 1943 BP.
 XX
 AC ABK13467;

XX 23-APR-2002 (first entry)
 DT DNA encoding Immunotoxin fusion protein DT389.sfv(UCHT1).
 DE
 XX Immunotoxin; pseudomonas exotoxin A; ETA; diphtheria toxin; DT;
 KW cytostatic; immunosuppressive; immunostimulant; antidiabetic;
 KW antirheumatic; antiarthritic; anti-HIV; anti-inflammatory;
 KW anti-T cell immunotoxin fusion protein; antibody; UCHT1;
 KW (Gly4Ser)3 linker; T cell leukaemia; lymphoma; gene; ds;
 KW graft-versus-host disease; Fv; autoimmune disease; transplant rejection;
 KW systemic lupus erythematosus; type I diabetes; rheumatoid arthritis;
 KW myasthenia gravis; multiple sclerosis; AIDS; DT389.sfv(UCHT1);
 KW acquired immunodeficiency syndrome; chronic immunosuppression.
 XX
 OS Mus sp.
 OS Corynebacterium diphtheriae.
 OS Synthetic.
 OS Chimeric.
 XX Key Location/Qualifiers
 CDS 3..1132
 /*tag= a
 /product= "DT389.sfv(UCHT1)"
 FT WO200187982-A2.
 FN 22-NOV-2001.
 PD 18-MAY-2001; 2001WO-US016125.
 PF 18-MAY-2000; 2000US-00573797.
 PR (NOVS) NOVARTIS AG.
 PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Neville DM, Thompson JT, Hu H, Woo J, Ma S, Hexham JM, Digan ME;
 XX WPI; 2002-121980/16.
 DR P-PSDB; AAU75382.
 DT
 XX New anti-T cell immunotoxin fusion protein comprising a truncated
 PT diphtheria toxin moiety, a connector, and one single chain Fv of the
 PT variable region of a UCHT1 antibody, useful for treating e.g. autoimmune
 PT diseases.
 XX
 XX Claim 34; Fig 35; 307pp; English.
 CC The invention relates to an anti-T cell immunotoxin fusion protein,
 CC comprising from the amino terminus, a truncated diphtheria toxin moiety,
 CC a connector, and one single chain Fv of the variable region of a UCHT1
 CC antibody. The single chain Fv comprises VL, L, or VH, L is a Gly-Ser
 CC linker, and VL and VH are the variable light and heavy domains of the
 CC anti-CD3 antibody UCHT1. Also included are a nucleic acid encoding the
 CC novel immunotoxin fusion protein, a vector comprising the nucleic acid
 CC and a cell comprising the nucleic acid. The immunotoxin may also be a
 CC pseudomonas exotoxin A (ETA). The immunotoxins are useful for treating T
 CC cell leukaemias or lymphomas, graft-versus-host diseases, and autoimmune
 CC diseases by inducing immune tolerance. The immunotoxin fusion proteins
 CC may be used in vivo to systemically reduce populations of T cells, or ex
 CC vivo to effect T-cell depletion from a treated cell population. The
 CC fusion proteins can be administered to a subject who is or will be a
 CC recipient of an allotransplant to prevent or reduce T-cell mediated acute
 CC or chronic transplant rejection of the transplanted allogeneic cells,
 CC tissue or organ in the subject as well as treat other T-cell mediated
 CC diseases such as systemic lupus erythematosus, type I diabetes,
 CC rheumatoid arthritis, myasthenia gravis, multiple sclerosis, infectious
 CC diseases of the immune system (e.g. AIDS (acquired immunodeficiency
 CC syndrome)) and chronic immunosuppression. The present sequence encodes an
 CC immunotoxin fusion protein of the invention comprising 389 residues from
 CC the N-terminal glycine of mature DT toxin, a (Gly4Ser)3 linker and one
 CC single Fv chain from antibody UCHT1, DT389.sfv(UCHT1)


```

Db 1675 GAGGATATGCCACTTACTTTTCCCAACAGGGTAATACGCTTCGCGGACGTTCCGCTGGA 1734
Qy 301 GGCACCAAGCTGGAATAAACAAGTGTGGAGTGGTCTGTGATGGTGGTGGTGGTGGTGGTGG 360
Db 1735 GGCACCAAGCTGAGATAAAGAGGCGGAGGAG-----CGGAGGC 1776
Qy 361 GCGGGCTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420
Db 1777 GGTGGTTCGGAGGGAGGAGGCTCGGAGGTGCAGCTCCAGAGTCTGGACCTGGAGTGGTGG 1836
Qy 421 AAGCCTGGGGCTCAGTGAAGATTTCTGCAAGATTTCTGGTACCGCATTCATAGCTCT 480
Db 1837 AAGCCTGGAGCTTCAATGAAGATATCTGCAAGGCTTCTGGTACTCAATCACTGGCTAC 1896
Qy 481 TGGATGAACCTGGTGAAGCAGAGGCTGGACAGGCTTGGATGGATTCGACGATTTAT 540
Db 1897 ACCATGAACCTGGTGAAGCAGAGTCTGGAAGAACCTTGGATGGATGGATTTAT 1956
Qy 541 CCTGGAGATGAGATTTAACTAACTGGGAATTCGAGGGCAAGGCCATACGACAGCA 600
Db 1957 CCTTACAAAGGTGTAGTACTCAACCAAGAGTTCGAAGCAAGGCCACATTAAGTGA 2016
Qy 601 GACAAATCTCCAGCAGAGCTACATGCAGCTCAGAGCTGACCTCTGGACTCTGG 660
Db 2017 GACAAGTCACTCCAGCAGAGCTACATGGAACCTCTCAGTCTGACATCTGAGGACTCTGCA 2076
Qy 661 GTCTATTCTGTGCAAGATCGGGTGTGCTAGCTTATGCTATGACTACTGGGCTCAAGGA 720
Db 2077 GTCTATTACTGTCAAGATCGGGTACTACGGTGTAGTACTGGTGTCTGATGTCGTGG 2136
Qy 721 ACCTCAGTCAAC 732
Db 2137 GCGCAGGCAAC 2148

RESULT 5
AAQ81076
ID AAQ81076 standard; cDNA; 913 BP.
AC
XX
XX
XX 25-MAR-2003 (revised)
DT 14-MAR-1995 (first entry)
XX
XX
DE Bispecific CD3-L6Fvlg antibody derivative coding sequence.
XX fusion protein; recombinant bispecific single chain antibody;
XX helical peptide linker; anti-L6 antibody; tumour cell antigen;
XX anti-CD3 antibody; variable region; ss.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
FH 7..912
FT /*tag= b
FT sig_peptide 7..75
FT /*tag= a
FT /*note= "L6 light chain variable region leader"
FT misc_RNA 76..819
FT /*tag= c
FT /*note= "encodes CD3 VL-VH fusion"
FT 820..828
FT /*tag= d
FT /*note= "encodes hinge region of Fc domain"
FT 829..912
FT /*tag= e
FT /*note= "encodes helical peptide linker"
XX
XX EP610046-A2.
XX
XX 10-AUG-1994.
XX
XX 31-JAN-1994; 94EP-00300692.

```

```

XX 01-FEB-1993; 93US-00013420.
PR 13-SEP-1993; 93US-00121054.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Ledbetter JA, Gilliland LK, Hayden MS, Linsley PS, Bajorath J;
XX Fell PH;
XX WPI; 1994-250885/31.
XX P-PSDB; AAR60206.
XX
XX Expression vector encoding bispecific fusion protein - having binding
XX domains for separate targets joined by helical peptide, useful e.g. for
XX diagnosis and treatment.
XX
XX Example 1; Fig 11; 50pp; English.
XX
XX The VL and VH sequences of the anti-CD3 hybridoma G19-4 were amplified by
XX PCR methods. A gene fusion was constructed from the two amplified domains
XX and a (Gly4Ser)3 linker. The amino terminus of the VL-VH fusion cassette
XX was fused at the SalI site to the L6 light chain variable region leader
XX peptide and the carboxy-terminus was fused directly to the hinge region
XX of the Fc domain at the BclI site and/or to a short "helical" peptide
XX linker to construct the bispecific CD3-L6Fvlg antibody derivative. The
XX variable regions for L6 were fused in frame to the opposite end of the
XX helical linker (not included in AAQ81076). (Updated on 25-MAR-2003 to
XX correct PN field.)
XX
XX Sequence 913 BP; 252 A; 223 C; 228 G; 210 T; 0 U; 0 Other;
XX
Query Match 64.7%; Score 479.2; DB 2; Length 913;
Best Local Similarity 79.4%; Pred. No. 1.7e-115;
Matches 581; Conservative 0; Mismatches 148; Indels 3; Gaps 1;
Qy 1 GATATTCAGATGACACAGACTACATCTCCCTCTGCTGCTCTCTGGGAGACAGAGTCAAC 60
Db 76 GACATCCAGATGACACAGACTACATCTCCCTCTGCTGCTCTCTGGGAGACAGAGTCAAC 135
Qy 61 GTCAGTTGAGGGCAAGTCAGGACATTAGGAAATTTAAATTTAAATTTAAATTTAAATTTAA 120
Db 136 ATCAGTTGAGGGCAAGTCAGGACATTCGCAATTTAAATTTAAATTTAAATTTAAATTTAA 195
Qy 121 GACGGAATGTTAAATTTCTGATCTACTACATCAAGATTTACTGCCAGGAGTCCCATCA 180
Db 196 GATGGAATGTTAAATTTCTGATCTACTACATCAAGATTTACTCGAGGAGTCCCATCA 255
Qy 181 AGGTTTCAGTGGCAGTGGTCTGGAAACAGATTTATTCCTCACATTTAAACAACTGGAGCAG 240
Db 256 AGGTTTCAGTGGCAGTGGTCTGGAAACAGATTTATTCCTCACATTTAAACAACTGGAGCAG 315
Qy 241 GAAGATTTGGCAGTCTACTTTTGGCAACAGGCAATAGCCCTCGGTGACGTTCCGTGGA 300
Db 316 GAAGATTTGGCAGTCTACTTTTGGCAACAGGCAATAGCCCTCGGTGACGTTCCGTGGA 375
Qy 301 GGCACCAAGCTGGAATAAAGCTGGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 360
Db 376 GGCACCAAGCTGGTAAACCAACGGGAGCTCGGTGGCGGTGGTGGCGGTGGTGGGTGG 435
Qy 361 GCGGCTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420
Db 436 GGTGG---CGGCGGATCTATCGATGAGGTCCAGCTGCAACAGTCTGGACCTGGAACCTGGTG 492
Qy 421 AAGCCTGGGGCTCAGTGAAGATTTCTGCAAGATTTCTGGCTAGCATTCATTAAGCTCT 480
Db 493 AAGCCTGGAGCTTCAATGACATGTCTTGCAGGCGCTCTGGTACTCATTCATTCGCTAC 552
Qy 481 TGGATGAACCTGGTGAAGCAGAGGCTGGACAGGCTTGGATGGATTCGACGAGGATTTAT 540
Db 553 ATCGTGAACCTGGTGAAGCAGAGGCTGGACAGGCTTGGATGGATTCGACGAGGATTTAT 612
Qy 541 CCTGGAGATGGAGATTTCTAACTACATGGAATTCGAGGGCAAGGCCATACGACAGCA 600

```


XX 26-MAR-2002; 2002US-00107991.
 XX (GENE-) GENE-CRAFT INC.
 XX Ledbetter JA, Hellstrom I, Hayden-Ledbetter M, Hellstrom KB;
 XX WPI; 2003-804045/75.
 XX DR P-PSDB; ADB86039.
 XX New culture system for generating tumor-reactive T-lymphocytes comprises
 PT T-cells from a cancer patient, antigen-presenting cells, autologous or
 PT allogeneic tumor cells, and immobilized antibodies to T-cell receptors.
 XX Example 2; SEQ ID NO 1; 112pp; English.

XX The present sequence is of a DNA construct encoding a mouse-human
 CC chimeric single chain (sc) Fv comprising the G19-4 mouse anti-human CD3
 CC VL domain joined via a peptide linker to the G19-4 mouse anti-human CD3
 CC VH domain, human IgG1 FC domain (hinge, CH2, CH3) and human CD80
 CC transmembrane and cytoplasmic domains, with an N-terminal L6 V kappa
 CC leader peptide. The coding sequence was constructed to provide expression
 CC of active anti-CD3 monoclonal antibody scFv at the tumour cell surface.
 CC The DNA construct, or a cell transfected with it, can be used for in vivo
 CC cancer therapy. It is believed that the anti-CD3 scFv expression on the
 CC surface of tumour cells induces polyclonal T cell activation and tumour
 CC cell destruction, releasing tumour antigens and promoting a transition to
 CC antigen-specific tumour immunity, detected as rejection of 'wild-type'
 CC (not transfected) cells from the same tumour. In an example from the
 CC invention, cells expressing anti-CD3 scFv at their surface were shown to
 CC induce polyclonal stimulation of T cells to proliferate, produce Th1 type
 CC lymphokines, to become cytolytic and to have antitumour activity in vivo.

XX SQ Sequence 1687 BP; 434 A; 470 C; 431 G; 352 T; 0 U; 0 Other;

Query Match 62.6%; Score 463.8; DB 9; Length 1687;
 Best Local Similarity 79.3%; Pred. No. 2.2e-111;
 Matches 594; Conservative 0; Mismatches 137; Indels 18; Gaps 3;

QY 1 GATATTGAGTACACAGATCATCTCCCTCTCTGCTCTCTGGAGACAGAGTCACC 60
 DB 76 GACATCCAGATGACACAGATCATCTCCCTCTCTGCTCTCTGGAGACAGAGTCACC 135
 QY 61 GTGAGTTCAGGCGCAAGTCAGGACATAGGATTAATTAATCTGATCAGCAGAAACCA 120
 DB 136 ATGAGTTCAGGCGCAAGTCAGGACATAGGATTAATTAATCTGATCAGCAGAAACCA 195
 QY 121 GACGGAATCTTAAATCTCTGATCTACTACATCAAGATTAATCTGCGAGGTCCTCCATCA 180
 DB 196 GATGGAACTGTTAAATCTCTGATCTACTACATCAAGATTAATCTGCGAGGTCCTCCATCA 255
 QY 181 AGGTTGAGTGGCAGTGGTCTGAAACAGATTAATTCCTCACCATTAAACCTGAGAGCAG 240
 DB 256 AGGTTGAGTGGCAGTGGTCTGAAACAGATTAATTCCTCACCATTAAACCTGAGAGCAG 315
 QY 241 GAAGATATTGGCACTTACTTTTCCCAACAGGCAATACGCTCCGTCGGAGCTTCGGTGA 300
 DB 316 GAAGATATTGGCACTTACTTTTCCCAACAGGCAATACGCTCCGTCGGAGCTTCGGTGA 375
 QY 301 GGACCAAGCTGGAATAAAGCTGGTGGAGGTTCTTCTGATGGTGGTGGTCTTGGCGGC 360
 DB 376 GGACCAAGCTGGAATAAAGCTGGTGGAGGTTCTTCTGATGGTGGTGGTCTTGGCGGC 435
 QY 361 GCGCGCTCCGTTGGTGGATCCGAGGTCCTCAACAGCTGCACTGAGCTGAGCTGGTG 420
 DB 436 GGTGG---CGCGGATCTATCGATAGGTCCTCAACAGCTGCACTGAGCTGAGCTGGTG 492
 QY 421 AAGCCTGGGCGCTCAGTGAAGATTTCTGCAAGATTTCTGGTACCGATTCATAGCTCT 480
 DB 493 AAGCCTGGGCGCTCAGTGAAGATTTCTGCAAGATTTCTGGTACCGATTCATAGCTCT 546
 QY 481 TGAATGAGTGGTGAAGCAGAGGCTTGCACAGGCTTTGAGTGGATTTGACGAGTTAT 540

DB 547 ATCGTGAAGTGGCTGAAGCAGAGCCATGGAAAGAACTTTCAGTGGATTGGACTTATTAT 606
 QY 541 CTTGGAGATGGAGATTCTTAATCAATGGGAAATTCGAGGCGCAAGCCATATCGACAGCA 600
 DB 607 CCATACAAAGGCTCTTACTACTACCAACCAAGAAATTCAGGCGCAAGCCATTAATCTGTA 666
 QY 601 GACAAATCTCCAGCAGACGCTACATGACCTCAGACGCTCAGACGCTGCTGGACTCTGCG 660
 DB 667 GACAAGTCATCCAGCAGACGCTACATGAGCTCTCTAGTCTGACATCTGAAGACTCTGCA 726
 QY 661 GTCTATTCTTGTGCAAGATCGGGTGTCTACTGTTATGC-----TATGGACTACTGG 711
 DB 727 GTCTATTACTGTGCAAGATCTGGGTACTATGGTGTACTCGGACTGGTACTTCGATGTCTGG 786
 QY 712 GGTCAAGAACCTCAGTCACGCTCTCTC 740
 DB 787 GGCACAGGACCCAGCGTCACCGTCTCTC 815

RESULT 8
 ABK13612
 ID ABK13612 standard; DNA; 2691 BP.
 XX
 AC ABK13612;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE DNA encoding an Immunotoxin fusion protein.
 XX
 XX Immunotoxin; pseudomonas exotoxin A; ETA; diphtheria toxin; DT;
 KW cytosolic; immunosuppressive; immunostimulant; antidiabetic;
 KW antirheumatic; antiarthritic; anti-HIV; anti-inflammatory;
 KW anti-T cell immunotoxin fusion protein; antibody; UCHL1;
 KW (Gly4Ser)3 linker; T cell leukaemia; lymphoma; gene; ds;
 KW graft-versus-host disease; Fv; autoimmune disease; transplant rejection;
 KW systemic lupus erythematosus; type I diabetes; rheumatoid arthritis;
 KW myasthenia gravis; multiple sclerosis; AIDS;
 KW acquired immunodeficiency syndrome; chronic immunosuppression.

XX Mus sp.
 OS Corynebacterium diphtheriae.
 OS Synthetic.
 OS Chimeric.
 XX WO200187982-A2.
 XX
 PD 22-NOV-2001.
 XX
 PF 18-MAY-2001; 2001WO-US016125.
 XX
 PR 18-MAY-2000; 2000US-00573797.
 XX
 PA (NOVS) NOVARTIS AG.
 PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Neville DM, Thompson JT, Hu H, Woo J, Ma S, Hexham JM, Digan ME;
 DR WPI; 2002-121980/16.
 XX
 XX New anti-T cell immunotoxin fusion protein comprising a truncated
 PT diphtheria toxin moiety, a connector, and one single chain Fv of the
 PT variable region of a UCHL1 antibody, useful for treating e.g. autoimmune
 PT diseases.
 XX
 PS Claim 91; Page 300-301; 307pp; English.
 CC
 CC The invention relates to an anti-T cell immunotoxin fusion protein,
 CC comprising from the amino terminus, a truncated diphtheria toxin moiety,
 CC a connector, and one single chain Fv of the variable region of a UCHL1
 CC antibody. The single chain Fv comprises VL, L, or VH, L is a Gly-Ser
 CC linker, and VL and VH are the variable light and heavy domains of the
 CC anti-CD3 antibody UCHL1. Also included are a nucleic acid encoding the

CC novel immunotoxin fusion protein, a vector comprising the nucleic acid
CC and a cell comprising the nucleic acid. The immunotoxin may also be a
CC pseudomonas endotoxin A (ETA). The immunotoxins may also be a
CC cell leukaemias or lymphomas, graft-versus-host diseases, and autoimmune
CC diseases by inducing immune tolerance. The immunotoxin fusion proteins
CC may be used in vivo to systemically reduce populations of T cells, or ex
CC vivo to effect T-cell depletion from a treated cell population. The
CC fusion proteins can be administered to a subject who is or will be a
CC recipient of an allotransplant to prevent or reduce T-cell mediated acute
CC or chronic transplant rejection of the transplanted allogeneic cells,
CC tissue or organ in the subject as well as treat other T-cell mediated
CC diseases such as systemic lupus erythematosus, type 1 diabetes,
CC rheumatoid arthritis, myasthenia gravis, multiple sclerosis, infectious
CC diseases of the immune system (e.g. AIDS (acquired immunodeficiency
CC syndrome)) and chronic immunosuppression. The present sequence encodes an
CC immunotoxin fusion protein of the invention comprising 390 residues from
CC the N-terminal glycine of mature DT toxin, a (Gly45Ser) 3 linker and one
CC single Fv chain from antibody UCHL1, and has had its AT content altered
CC to enable expression in yeast

Sequence 2691 BP: 669 A; 649 C; 701 G; 672 T; 0 U; 0 Other;

Query Match 58.3%; Score 431.8; DB 6; Length 2691;
Best Local Similarity 76.1%; Pred. No. 6.1e-103;
Matches 570; Conservative 0; Mismatches 152; Indels 27;

Qy	1	GATATT	CAGATG	CACAG	CTAC	TAC	CTCT	CCTG	TCTGC	CTCT	CTGG	GAG	CAG	GTCA	CC	60
Db	1180	GACATC	CAGATG	ACCC	AG	CA	CA	CTCT	CCTG	TCTGC	CTCT	CTGG	CG	CAG	GTCA	CC
Qy	61	GTCA	TTG	CAGG	CA	GT	CA	GG	CA	TTAG	GA	ATTA	TTAA	CT	GG	TA
Db	1240	ATCA	TTG	CAGG	CA	GT	CA	GG	CA	TTAG	GA	ATTA	TTAA	CT	GG	TA
Qy	121	GACG	GA	ACTG	TTAA	TT	CT	GA	TCTA	CA	CA	TCA	AG	ATTA	CT	GC
Db	1300	GACG	GA	ACTG	TTAA	TT	CT	GA	TCTA	CA	CA	TCA	AG	ATTA	CT	GC
Qy	181	AG	TTT	CAG	TG	CG	AG	TG	CG	AG	TTAT	TTC	CT	CA	CA	CT
Db	1360	AG	TTT	CAG	TG	CG	AG	TG	CG	AG	TTAT	TTC	CT	CA	CA	CT
Qy	241	GA	GA	TAT	TG	CA	CTTA	CT	TTT	TG	CA	CA	G	GG	CA	AT
Db	1420	GAG	GA	TAT	TG	CA	CTTA	CT	TTT	TG	CA	CA	G	GG	CA	AT
Qy	301	GG	CA	CA	AG	CT	GG	AA	ATA	AA	AC	GT	GG	TA	AG	GT
Db	1480	GG	CA	CA	AG	CT	GG	AA	ATA	AA	AC	GT	GG	TA	AG	GT
Qy	361	GG	CG	CT	CC	GT	GT	GT	GG	AT	CC	AG	GT	CC	AA	CT
Db	1522	GG	CG	CT	CC	GT	GT	GT	GG	AT	CC	AG	GT	CC	AA	CT
Qy	421	AAG	CT	GG	GG	CT	CAG	TG	AG	AT	TC	CT	GC	AA	GA	AT
Db	1582	AAG	CT	GG	GG	CT	CAG	TG	AG	AT	TC	CT	GC	AA	GA	AT
Qy	481	TG	GA	TAA	CT	GG	GT	GA	AG	CA	GG	CT	CG	CA	GG	AT
Db	1642	ACC	AT	GA	CT	GG	GT	GA	AG	CA	GG	CT	CG	CA	GG	AT
Qy	541	CCT	GG	AG	AT	GG	AA	TT	CT	CA	AT	GG	GA	AA	TT	CG
Db	1702	CCT	TTA	CA	AG	GG	GT	CT	CG	AC	TA	CA	CC	AG	AA	TT
Qy	601	GAC	AA	AT	CCT	CG	AG	CA	GG	CT	TA	CT	GC	AG	CT	GC
Db	1762	GAC	AA	AT	CCT	CG	AG	CA	GG	CT	TA	CT	GC	AG	CT	GC
Qy	661	GT	CT	AT	TT	CT	GT	CA	AG	AT	CG	GT	TT	GC	TA	CT
Db	1822	GT	CT	AT	TT	CT	GT	CA	AG	AT	CG	GT	TT	GC	TA	CT

Qy 712 GGTAAGGAACCTCAGTACCGTCTCTC 740
Db 1882 GGTGCTGGTACTACTGTCACTGTCTCTC 1910

RESULT 9

ACF05482 standard: cDNA: 1953 bp.

AC ACF05482:

DT 06-NOV-2003 (first entry)

XX Anti-CD7 antibody TH-69 scFv-Pseudomonas exotoxin A fusion sequence.

AA CD7; antibody; TH-69; scFv; exotoxin A; immunotoxin; cytostatic;
KW immunosuppressive; vaccine; gene; ss.

XX Mus sp.

OS Mus sp.
OS Pseudomonas sp.
OS Synthetic.

33 XX

XX	Key	Location/Qualifiers
FH		

FT	CDS	1. .1953
----	-----	----------

```
FT
/*tag= a
```

```
FT /partial
FT /product= "TH-69 scFv-ETA fusion protein"
```

PN WO2003051926-A2.

26-JUN-2003

11-DEC-2002: 2002WO-EP014064.

14-DEC-2001: 2001US-0339422P.

XX PA (UYER-) UNIV ERLANGEN-NUERNBERG.

XX
PT
Fev GHM. Gramatzki M. Peipp M:

XX
DR WPI: 2003-523519/49.

DR P-PSDB: ABR62591;

New fusion protein comprising functionally linked components of an anti-CD7 antibody or its fragment and an immunotoxin, useful for treating CD7 positive acute T-cell and/or myeloid leukemias, or Graft-versus-host disease.

XX
PS
Claim 11: page 10-11: 55pp: English.

The present sequence encodes a novel fusion protein comprising an scFv fragment of anti-CD7 monoclonal antibody TH-69 and domains II and III of Pseudomonas exotoxin A (ETA), joined via a peptide linker, and including an N-terminal 6xHis tag and C-terminal KDEL sequence. This novel fusion protein binds to CD7-positive T-lymphoid cells and kills them by the induction of apoptosis. It is an example of anti-CD7 scFv immunotoxin fusion proteins of the invention. An expression vector, host cells, a method for producing the fusion protein, and a vaccine comprising the fusion protein are claimed. The fusion protein is used in the treatment of disorders involving a hyperproliferation of CD7-positive cells, especially acute T-cell and/or myeloid leukaemia, and also in the treatment or prophylaxis of Graft-versus-host disease. A claimed method of conditioning an animal (e.g. a human) to be transplanted with donor cells, tissue or organ comprises: (a) depleting the CD7-bearing cell population in the animal; (b) providing a transplant comprising isolated bone marrow and/or stem cell-enriched peripheral blood cells of the donor, where the CD7-bearing cell population in (a) and/or the transplant in (b) are treated with a T-cell depleting effective amount of the fusion protein or therapeutic composition comprising the fusion protein; and (c) introducing the transplant into the animal.

QY	181	AGGTTTCAGTGGCAGTGGGCTCTGGAAACAGATATTTCCCTCACCATTAAACAACCTGGAGCAG	240
DB	229	AGGTTTCAGTGGCAGTGGGCTCTGGAAACAGATATTTCTCTCACCATTACAGAACTCTGAACT	288
QY	241	GAAGATATTTGGCAGTCTACTTTTGGCAACAGGGCAATACGCTCCGTGGACGTTCTGGTGGAA	300
DB	289	GAAGATATTTGCCACTTATTTATTTGTCAGCAGTATAGCAAGCTTCCGTACACGTTCCGAGGG	348
QY	301	GGCACCAAGCTGGAAATTAACACGTTGGAGGTGGTTCTGATGGTGGTGGTCTGGCGGC	360
DB	349	GGGACCAAGCTGGAAATTAACACGTTGGAGGTGGTTCTGATGGTGGTGGTCTGGCGGC	408
QY	361	GGCGGCTCCGCTGGTGGTGGATCCGAGGTCCAACTGCACAGCTCTGGACCTGAGCTGGTG	420
DB	409	GGCGGCTCCGCTGGTGGTGGATCCGAGGTGCAACTGGTGGAGTCTGGGGAGGCTTAGTG	468
QY	421	AGGCTGGGGCCCTCAGTGAAGATTTCTGCAAGATTTCTGGCTACGCATTTCAATAGCTCT	480
DB	469	AGGCTGGGGGGTCCCTGAACTCTCTCTGTCAGCCTCTGCACTCACTTTTCAGTAGCTAT	528
QY	481	TGATGAACTGGGCTGAAGCAGAGGGCTCTGGACAGGGTCTTGAGTGGATTTGACGGATTAT	540
DB	529	GCCATGTCTGGGTTCCGCCAGACTCCAGAGAAGAGGCTGGAGTGGTGGCATCCATTAGT	588
QY	541	CCTGGAGATGGAGATTTCTAACTACAAATGGGAAATTCGAGGGCAAGGCCATACATGACAGCA	600
DB	589	AGTGG---TGGTTTCACTACTATCCACAGCAGTGTGAAGGGCCGATTCACCATCTCCAGA	645
QY	601	GACAAATCCTCCAGCACAGCCTACATGACGCTCAGCAGCCTGACCTCTGTGGACTCTGGCG	660
DB	646	GATAATGCCAGGAACATCCGTATCTGCAAAATGAGCAGTCTGAGGTCTGAGGACACGGCC	705
QY	661	GTCTATTCTGTGCGAGATCGGGGTTGCTAGCTATGCTATGAGTACTTGGGCTCAAGGA	720
DB	706	ATGTATTACTGTGCAAGACAGAGTACGGGGGTA---CCTCGATGTCTGGGGCGCAGGG	762
QY	721	ACCTCAGTCAACGCTCTCCCTCG	741
DB	763	ACCAAGTCAACGCTCTCCCTCG	783
RESULT 11			
AAA39167			
ID	AAA39167 standard; DNA; 1457 BP.		
XX	AAA39167;		
XX	05-SEP-2000 (first entry)		
DT	Apoptobody3sc fusion protein encoding nucleotide sequence SEQ ID NO:7.		
XX	Human; Fas antigen; apoptosis; apoptobody3sc; antibody; fusion gene;		
KW	medical; pharmaceutical; pharmacological; biochemical; db.		
OS	Homo sapiens.		
OS	Synthetic.		
FH	Key		
FT	CD5		
FT	Location/Qualifiers		
FT	7,1317		
FT	/*tag= a		
FT	/product= "apoptobody3sc"		
XX	JP2000102389-A.		
PN	11-APR-2000.		
PD	29-SEP-1998; 98JP-00291441.		
PF	29-SEP-1998; 98JP-00291441.		
XX	(HAGI/) HAGIWARA Y.		
PA	(HAGI/) HAGIWARA H.		
XX			

OS	Synthetic.
XX	DE19819846-A1.
FN	
XX	11-NOV-1999.
PD	
XX	
PF	05-MAY-1998; 98DE-01019846.
XX	
PP	05-MAY-1998; 98DE-01019846.
PR	(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX	
PA	Little M, Kipriyanov S;
XX	
PI	WPI; 2000-024472/03.
XX	P-PSDB; AAY50822.
DR	Multivalent Fv-antibody constructs with at least four variable domains
DR	connected by 1, 2 and 3 peptide linkers.
PT	Example 1; Fig 5; 14pp; German.
XX	
PS	This invention describes a novel multivalent Fv-antibody construct with
XX	at least four variable domains that are connected to one another by 1, 2
CC	and 3 peptide linkers. The construct has antiviral, antibacterial and
CC	cytostatic activity. The multivalent Fv-antibody constructs are useful
CC	for the diagnosis and/or therapy of disease, especially viral, bacterial
CC	or tumor diseases. The multivalent Fv-antibody constructs have increased
CC	stability when in the form of a single chain dimer. This sequence encodes
CC	a bivalent Fv antibody construct composed of the antibody 9E10 epitope in
CC	expression plasmid pDISC3x19-LI
XX	
QQ	Sequence 1698 BP; 433 A; 434 C; 450 G; 381 T; 0 U; 0 Other;
Query Match	54.5%; Score 404.2; DB 3; Length 1698;
Best Local Similarity	73.5%; Pred. No. 9e-96;
Matches	577; Conservative 0; Mismatches 163; Indels 45; Gaps 3;
QY	1 GATATTGAGATGACACAGACTACATCTCCCTGTCTGCCTCTCTGGGAGACAGAGTCACC 60
Db	478 GATATCTTGGCTCACCAAACTCCAGCTTCTTTGGCTGTGCTCTAGGGCAGAGGCCACC 537
QY	61 GTCAATTGCGAGGCAAGTCAG-----GACATTAGGAATTAATTTAAACTGGTAT 108
Db	538 ATCTCTCGAAGGCCAGCCAAGTGTTGATATGATGGTGATAGTATTATTGAACCTGGTAC 597
QY	109 CAGCAGAAAACAGACGGAACGTGTTAAATTCCTCATCTACTACATCAAGATTACTGCGCA 168
Db	598 CAACAGATTCCAGGACAGCACCACAACTCCTCATCTATGATGCATCCAATCTAGTTCT 657
QY	169 GGAGTCCCATCAGGTTTCGTGGCAGTGGCTCTGGACAGATTATTCCTCCACCAATTAC 228
Db	658 GGGATCCCACCCAGTTTTAGTGGCAGTGGGCTCTGGGACAGACTTCACTCCCTCAACATCCAT 717
QY	229 AACCTGGAGCAGGAAGATATTGGCAGCTTACTTTTGCCAACAGGGCAATACGCTCCGTGG 288
Db	718 CCTGTGGAGAAGGTGGATGATGCAACCTATCATCTGTGAGAAAGTACTGAGGATCCGTGG 777
QY	289 ACCTTCGGTGGAGCACCACAGCTGGAATAAAG-----TGGTGGA 330
Db	778 ACCTTCGGTGGAGCACAAGCTGGAATCAAACGGGCTGATGCTGGCGCGCTGGTGGT 837
QY	331 GGTGGTTCTGATGGTGGTGGTTCTGGCGCGGGGCTCCGGTGGTGGTAGATCCGAGGTC 390
Db	838 GGTGGTTCTGGCGCGGCTGGTAGCGGTGGTGGCGGCTCCGGTGGTGGTAGCCAGGTG 897
QY	391 CAACCTGCAACAGCTCTGGACCTGAGCTGGTGAAGCTCGGGGCTCAGTGAAGATTTCCTGC 450
Db	898 CAGCTGCAGAGTCTGGGGCTGAGCTGGTAGGGCTGGGTCTCTCAGTGAAGATTTCCTGC 957
QY	451 AAAGATTCTGGCTACGATTAATAGCTCTTGGATGAACCTGGGTGAAGCAGAGCCCTGGA 510
Db	958 AAGGCTTCTGGCTATGCAATCAGTAGCTTACTGGATGAATCCTGGTGAAGCAGAGCCCTGGA 1017

Qy	511	CAGGGCTTCGAGTGGATGGACGGATTATTCCTGGAGATGGAGATTCTTAACCTACAATGGG	570
Db	1018	CAGGGCTTCGAGTGGATGGACAGATTTGGCCCTGGAGATGGTGATACTAACTACAATGGA	1077
Qy	571	AAATTTCAGGGCAAGCCCATCTACGACGACAGACAAATTCCTCCASCACAGCCTACATGCAG	630
Db	1078	AAGTTCAGGGTAAAGCCACTCTGACTGCGACAGCAATCCTCCAGCACAGCCTACATGCAA	1137
Qy	631	CTCAGCAGCCTGACCTCTGTGGACTCTGGGCTCTATTTCTGTGCAAGATCG-----	681
Db	1138	CTCAGCAGCCTGACCTCTGTGGACTCTGGGCTCTATTTCTGTGCAAGCGGAGACTACG	1197
Qy	682	-----GGGTTGCTACGTTATGCTATGCTATGGACTACTGGGGTCAAGAAACCTCAGTCACCGTC	735
Db	1198	ACGGTAGGGCGTTATTACTATGCTATGGACTACTGGGGTCAAGAAACCTCAGTCACCGTC	1257
Qy	736	TCCTC 740	
Db	1258	TCCTC 1262	
RESULT 14			
ACC79606			
ID	ACC79606	standard; DNA; 1817 BP.	
XX	ACC79606;		
AC	ACC79606;		
DT	05-AUG-2003	(first entry)	
XX			
DE	Plasmid pSKK2 scFv3-LL-Db19	nucleotide sequence.	
XX			
KW	Multimeric single chain tandem Fv-antibody; antibacterial; virucide;		
KW	cytostatic; cytokine antagonist; diagnosis; viral disease; prion disease;		
KW	bacterial disease; tumoral disease; agglutination; red blood cell;		
XX	immune system; tumour cell; cytokine; cytotoxic; gene; ds.		
OS	Synthetic.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	123..1811	
FT		/*tag= a	
FT		/product= "Plasmid pSKK2 scFv3-LL-Db19 protein"	
XX			
FN	EPI293514-A1.		
XX			
PD	19-MAR-2003.		
XX			
PF	14-SEP-2001; 2001EP-00122104.		
XX			
PR	14-SEP-2001; 2001EP-00122104.		
XX			
PA	(APFI-) AFFIMED THERAPEUTICS AG.		
XX			
PI	Le Gail F, Kipriyanov S, Reusch U, Moldenhauer G, Little M;		
XX			
DR	WPI; 2003-395536/38.		
DR	P-PSDB; ABR57058.		
XX			
PT	New multimeric Fv-antibody having monomers forming antigen-binding units		
PT	and sites, useful for the preparation of a pharmaceutical composition for		
PT	the treatment of a viral, bacterial, tumoral or prion diseases.		
XX			
PS	Example 1; Fig 6A-B; 29pp; English.		
XX			
CC	The present invention describes a multimeric Fv-antibody (I) comprising:		
CC	(a) monomers comprising at least 4 variable domains of which two		
CC	neighbouring domains of one monomer form an antigen-binding VH-VL or VL-		
CC	VH scFv unit; and/or (b) at least 2 variable domains of a monomer that		
CC	are non-covalently bound to 2 variable domains of another monomer		
CC	resulting in the formation of at least 2 additional antigen binding sites		
CC	to form multimerisation motif. Also described is a process for the		
CC	preparation of (I) comprising ligating DNA sequences encoding the peptide		

CC linkers with the DNA sequences encoding the variable domains such that
CC the peptide linkers connect the variable domains resulting in the
CC formation of a DNA sequence encoding a monomer of the multivalent
CC multimeric Fv-antibody, and expressing the DNA sequences encoding the
CC various monomers in an expression system. (I) has antibacterial, virucide
CC and cytostatic activities, and can be used as a cytokine antagonist. The
CC multimeric Fv-antibody is useful for diagnosis. The antibody can also be
CC used for the preparation of a pharmaceutical composition for the
CC treatment of a viral, bacterial, tumoural or prion disease, the
CC agglutination of red blood cells, linking cytotoxic cells of the immune
CC system to tumour cells, or linking activating cytokines, cytotoxic
CC substances or a telomerase to a target cell. The present sequence
CC represents the plasmid pSK2 sefv3-IL-Db19 nucleotide sequence, which is
CC used in the exemplification of the present invention
XX
XX Sequence 1817 BP; 464 A; 457 C; 480 G; 416 T; 0 U; 0 Other;
Query Match 53.4%; Score 395.8; DB 7; Length 1817;
Best Local Similarity 73.1%; Pred. No. 1.5e-93;
Matches 565; Conservative 0; Mismatches 172; Indels 36; Gaps 3;
QY 1 GATATTCAGATGACACAGACTACATCTCCCTGCTCTGCTGCTGCGGACAGAGTCCAC 60
Db |||||
QY 597 GATATCGTGTCTACTGATCTCCAGCAATCATGCTCTGCATCTCCAGGGGAGAGTCCAC 656
Db |||||
QY 61 GTGAGTTGAGGGCAAGTCAGACATTAGGAATTTAACTGTGTATCAGCAGAAACCA 120
Db |||||
QY 657 ATGACCTGCAGTCCAGCTCAAGTGAAG---TTACATGAACCTGTGTACAGCAGAGTCA 713
Db |||||
QY 121 GAGGAACTGTTAAATTCCTGATCTACTACACATCAAGATTACTGCAGAGAGTCCCATCA 180
Db |||||
QY 714 GGACCTCCCCAAGATGATTTATGACACATCCAACTGGCTTCTGGAGTCCCTGTGT 773
Db |||||
QY 181 AGGTTGAGTGGGAGTGGTCTGAAACAGATTATTCCTCTCACCATTAACTGAGGACAG 240
Db |||||
QY 774 CACTTCAGGGGCAAGTGGTCTGGGACCTCTTACTCTCTCACAATCAGCGCATGGAGGT 833
Db |||||
QY 241 GAGATATTGGCACTTACTTTTCCACAGGGCAATACGCTCCGTTGGAGCTTCGTGGA 300
Db |||||
QY 834 GAGATGCTGCCACTTATTACTGCCAGCATGGAGTAGTAACTTACGTTTCGGCTCG 893
Db |||||
QY 301 GGCAACCAAGCTGGAATAAATCAACG-----TGTGGAGGTGGTCTTGAT 342
Db |||||
QY 894 GGGACAAAGTTGGAATAAATCAACCGGCTGATCTCGGCGCGCTGGTGGTGGTCTGCG 953
Db |||||
QY 343 GGTGGTGGTCTGCGCGCGCGGCTCGGTGGTGGTGGATCCGAGTCCCACTGCAACAG 402
Db |||||
QY 954 GCGGCTGGTAGCGGTGGTGGCGGCTCGGTGGTGGTGGTAGCCAGGTGCGAGCAG 1013
Db |||||
QY 403 TCTGGACCTGAGCTGGTGGAGCCTGGGGCCTCAGTGAAGATTTCCTGCAAGATTCTGGC 462
Db |||||
QY 1014 TCTGGGCTGAGCTGGTGGAGCCTGGGTCTCAGTGAAGATTTCCTGCAAGGTCTTGGC 1073
Db |||||
QY 463 TACGATTCATAGCTCTTGGATGAATGGGTGAAGCAGAGGCTGGACAGGCTCTTGAG 522
Db |||||
QY 1074 TATGATTCAGTAGTACTAGTGAATGGGTGAAGCAGAGGCTGGACAGGCTCTTGAG 1133
Db |||||
QY 523 TGGATTGAGCGGATTTATCTGGAGATGGAGATTCTAACTACATGGGAATTCGAGGGC 582
Db |||||
QY 1134 TGGATTGAGCAGATTGGCTTGGAGATGGTGAATCTAACTAATGGAAGTTCAGGCT 1193
Db |||||
QY 583 AAGGCCATATGACAGCAGCAAAATCTCCAGCAGACGCTTACATGATGAGCTCAGAGCCTG 642
Db |||||
QY 1194 AAGGCCATCTGACTGACAGCAAAATCTCCAGCAGACGCTTACATGATGAGCTCAGAGCCTA 1253
Db |||||
QY 643 ACCCTGTGGACTCTGGGTCTATTCTGTGCAAGATCG-----GGTTG 687
Db |||||
QY 1254 GCATCTGAGGACTCTGGGTCTATTCTGTGCAAGATCGGAGACTAGCAGGTAGGCGGT 1313
Db |||||
QY 688 CTAGCTTATGCTATGAGTACTCTGGGGTCAAGGAACTTCAGTACCGCTCTCTC 740
Db |||||
QY 1314 TATTACTATGCTATGAGTACTCTGGGGTCAAGGAACTTCAGTACCGCTCTCTC 1366
Db |||||

RESULT 15

RAZ30332
ID AA230332 standard; cDNA; 1611 BP.
XX
AC AA230332;
XX
DT 11-FEB-2000 (first entry)
XX
DE Nucleotide sequence of the bscCD19xCD3 antibody.
XX
KW bscCD19xCD3 antibody; bispecific single-chain fragment; CD19 antigen;
KW CD3 antigen; CD19-positive target cell; T-cell stimulation;
KW cytotoxic T-lymphocyte; B-cell malignancy; myasthenia gravis;
KW B-cell mediated autoimmune disease; Morbus Basedow;
KW Hashimoto thyroiditis; Goodpasture syndrome; B-cell depletion;
KW non-Hodgkin lymphoma; gene therapy; cancer; viral disease; ds.
XX
OS Synthetic.
FH Key Location/Qualifiers
XX CDS 11..1606
FT /*tag= a
FT /*product= "bscCD19xCD3 antibody"
FT sig_peptide 11..67
FT /*tag= b
FT misc_feature 92..424
FT /*tag= c
FT misc_feature 470..841
FT /*tag= d
FT misc_feature 857..1213
FT /*tag= e
FT /*note= "encodes VH-CD3"
FT misc_feature 1268..1585
FT /*tag= f
FT misc_feature 1586..1603
FT /*tag= g
FT /*note= "encodes a His tag"
XX WO9954440-A1.
XX
XX 28-OCT-1999.
XX
XX 21-APR-1999; 99WO-EP002693.
XX
XX 21-APR-1999; 98EP-00107269.
XX (DOER/) DOERKEN B.
XX (RIET/) RIETHMUELLER G.
XX
XX Kufer P, Lutterbuese R, Bargou R, Loeffler A;
XX
XX WPI; 2000-013241/01.
XX P-PSDB; AAY43749.
XX
XX Novel multifunctional polypeptide for treating B-cell malignancies
XX especially non-Hodgkin lymphoma.
XX
XX Example 2; Fig 8; 91pp; English.
XX
XX The present sequence encodes a bscCD19xCD3 antibody. This antibody is a
XX bispecific single-chain polypeptide comprising domains providing binding-
XX site of immunoglobulin chains or antibodies specifically recognizing CD19
XX and CD3 antigen. The polypeptide destroys CD19-positive target cells
XX without any need of T-cell pre and/or co-stimulation, by recruiting
XX cytotoxic T-lymphocytes and so specific lysis by T-cells rather than a
XX direct effect by an antibody is achieved. The bispecific single-chain
XX polypeptides, or nucleotides encoding them, are used for the treatment of
XX B-cell malignancies, B-cell mediated autoimmune diseases like myasthenia
XX gravis, Morbus Basedow, Hashimoto thyroiditis or Goodpasture syndrome or

CC for the depletion of B- cells and more particularly non-Hodgkin lymphoma
CC in mammals preferably human. They can also delay the pathological
CC conditions caused by these diseases, and can be used for detecting these
CC diseases. The polynucleotide is used for gene therapy. The polypeptides
CC are also used for identifying compounds modulating B-cell/T-cell mediated
CC immune response with can in turn be used for treating cancer, its related
CC diseases and also for inhibiting viral diseases by preventing viral
CC infection

XX Sequence 1611 BP; 402 A; 396 C; 440 G; 373 T; 0 U; 0 Other;

```
Query Match          53.2%; Score 394.2; DB 3; Length 1611;
Best Local Similarity 73.5%; Pred. No. 3.7e-93;
Matches 564; Conservative 0; Mismatches 158; Indels 45; Gaps 3;

QY 1 GATATTCAGATGACAGACTACATCCCTCTGCTGCTCTCTGGGACACAGATCACC 60
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GTCAGTTGACAGGCAAGTCCAG-----GACATTAGGAATATTAACTGGTAT 108
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 152 ATCTCTGCAAGCCAGCCAAAGTGTGATTATGATGGTGATGTTATTTGAACGGTAC 211
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 109 CAGCAGAAACACAGCGAATGTTAAATCCCTGATCTACTACACATCAAGATTACTGCCA 168
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 212 CAACAGATTCACGAGACAGCCACCAACTCCTCATCTATGATGCATCCAACTAGTTCT 271
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 169 GGAGTCCCATCAAGGTTTCAGTGGCAGTGGTCTGGAACAGATTATTCCTCACCATTAC 228
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 272 GGGATCCACCCAGGTTTATGTCAGTGGTCTGGGACAGACTTCACCCCTCAACATCCAT 331
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 229 AACCTGAGCAGAGAGATATGGACTTACTTTTGGCAACAGGGCAATAGCCTCCGTGG 288
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 332 CCTGTGGAGAAGTGGATGTCGAACCTATCACTGTGACAAAGTACTAGGATCCGTGG 391
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 289 ACGTTCCGTGGAGCACCAAGCTGGAATAAAGCTGGTGGAGGTGGTTCTGATGGTGGT 348
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 392 ACGTTCGTGGAGGACCAAGCTCGATCAAA-----GGTGGTGGT 433
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 349 GGTTCGCGCGCGCGCTCCGGTGGTGGATCCGAGGTCCAACTGCMAAGTCTGGA 408
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 434 GGTTCGCGCGCGCGCTCCGGTGGTGGTGGTCTCAGGTGCAGCTGCAGCAGTCTGGG 493
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 409 CCTGAGCTGGTGAAGCCTGGGGCTCAGTGAAGATTTCCTGCAAGATTCTGGCTACGCA 468
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 494 GCTGAGCTGGTGAAGCCTGGGGCTCAGTGAAGATTTCCTGCAAGATTCTGGCTACGCA 553
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 469 TTCATAGCTCTTGGATGAATGGGTGAAGCAGAGGCTCGAAGGCTCTTGAAGTGGATT 528
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 554 TTCAGTAGCTACTGGATGAATGGGTGAAGCAGAGGCTCGAAGGCTCTTGAAGTGGATT 613
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 529 GGACGGATTTATCCTGGAGTGGAGATTCTAACTACAAATGGAAATTCGAGGGCAAGGCC 588
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 614 GGACAGATTTGGCTGGAGATGGTGAATCTAACTAATGAAAGTTCAAGGTTAAGCC 673
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 589 ATACTGACAGCAGACAAATCCTCAGACAGCCCTACATGAGCTCAGACGCTGACCTCT 648
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 674 ACTCTGACTGACAGCAATCCTCAGACAGCCCTACATGCAACTCAGACGCTAGCATCT 733
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 649 GTGACTCTGGGTCTATTCTGTGCAAGTCG-----GGGTTGCTACGT 693
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 734 GAGGACTCTGGGTCTATTCTGTGCAAGCGGAGACTAGCAGGTAGGCCGCTTATTAC 793
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 694 TATGCTATGACTACTGGGTCAAGGAACCTCAGTACCCTCTCCTC 740
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 794 TATGCTATGACTACTGGGCAAGGACCAACGCTCACCCTCTCCTC 840
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: April 26, 2004, 16:01:08
Job time : 399 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 15:43:37 ; Search time 3232 Seconds
(without alignments)
9937.249 Million cell updates/sec

Title: US-10-620-049-24

Perfect score: 741

Sequence: 1 gatattcagatgacacagac.....cttcagtcacgctctctcgcg 741

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg_hum.*

31: em.htg_inv.*

32: em.htg_other.*

33: em.htg_mus.*

34: em.htg_pln.*

35: em.htg_rod.*

36: em.htg_mam.*

37: em.htg_vit.*

38: em.sy.*

39: em.htgo_hum.*

40: em.htgo_mus.*

41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	499.4	67.4	1803	6	BD237850 Anti-CD3
2	499.4	67.4	1940	6	AX306601 Sequence
3	499.4	67.4	1943	6	AX306600 Sequence
4	482.8	65.2	2952	6	AX306590 Sequence
5	479.2	64.7	916	6	AR116105 Sequence
6	479.2	64.7	916	6	I45604 Sequence 29
7	479.2	64.7	916	6	AR256099 Sequence
8	479.2	64.7	916	6	AR401482 Sequence
9	456.2	61.6	748	12	AF326561 Synthetic
10	431.8	58.3	2691	6	AX306662 Sequence
11	421	56.8	1953	6	AX797328 Sequence
12	421	56.8	1956	6	AX797327 Sequence
13	418.6	56.5	1457	6	E35160 Method for
14	406	54.8	1545	6	AX398258 Sequence
15	404.2	54.5	1698	6	AX011206 Sequence
16	404.2	54.5	1698	6	BD206134 Polyvalen
17	396	53.4	888	12	AF132308 Synthetic
18	395.8	53.4	1817	6	AX739885 Sequence
19	394.2	53.2	1611	6	AX014270 Sequence
20	394.2	53.2	1611	6	AX839766 Sequence
21	394.2	53.2	1611	6	BD205046 CD19XCD3-
22	393	53.0	902	12	XXU49832 Synthetic s
23	385.6	52.0	1637	6	AX006750 Sequence
24	385.6	52.0	1637	6	BD222155 Immunolog
25	381.8	51.5	819	6	AI8692 Synthetic n
26	374.8	50.6	1817	6	AX739887 Sequence
27	371.2	50.1	1612	6	AX057943 Sequence
28	368	49.7	786	6	AX798501 Sequence
29	368	49.7	993	6	AX798489 Sequence
30	368	49.7	1017	6	AX798479 Sequence
31	363.8	49.1	2871	6	AX023363 Sequence
32	363.8	49.1	2871	6	BD222937 Heteromin
33	361.8	48.8	1653	6	AX011208 Sequence
34	361.8	48.8	1653	6	BD206135 Polyvalen
35	358.6	48.4	877	6	E31225 Device for
36	358.6	48.4	925	6	E30616 Antibody an
37	356.8	48.2	840	12	AF329458 Synthetic
38	356.4	48.1	1299	6	AR060677 Sequence
39	356.4	48.1	1299	6	AR074434 Sequence
40	356.4	48.1	1320	6	AR060674 Sequence
41	356.4	48.1	1320	6	AR074431 Sequence
42	356.4	48.1	6727	6	AR060673 Sequence
43	356.4	48.1	6727	6	AR074430 Sequence
44	356.4	48.1	6799	6	AR060676 Sequence
45	356.4	48.1	6799	6	AR074433 Sequence

ALIGNMENTS

RESULT 1
BD237850
LOCUS BD237850 1803 bp DNA linear PAT 17-JUL-2003
DEFINITION Anti-CD3 immunotoxins and therapeutic uses therefor.
ACCESSION BD237850
VERSION BD237850.1 GI:33047620
KEYWORDS JP 2002534441-A/1.
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 1803)
AUTHORS Digan,M.E., Lake,P. and Wright,R.M.
TITLE Anti-CD3 immunotoxins and therapeutic uses therefor
JOURNAL Patent: JP 2002534441-A 1 15-OCT-2002;
NOVARTIS AG

		OS	Artificial Sequence		
COMMENT		PJ	JP 2002534441-A/1		
		PD	15-OCT-2002		
		PF	13-JAN-2000 JP 2000593098		
		PR	15-JAN-1999 US 09/232445, 25-JAN-1999 US 09/236968 PR		
		PI	MARY ELLEN DIGAN, PHILIP LAKE, RICHARD MICHAEL WRIGHT PC		
		A6IK39/395,A6IK31/137,A6IK31/27,A6IK31/343,A6IK31/42,A6IK31/PC			
		PC	A6IK31/52,		
		A6IK31/573,A6IK31/664,A6IK31/7056,A6IK38/00,A6IK39/104 PC			
		A6IK37/06,C07K14/21,			
		C07D43/28,C07K19/00,C12N15/09//C07D261/18,C07D307/88 PC			
		D1937/38,C07D498/18,			
		C12N15/00,A6IK37/02			
		Description of Artificial Sequence: scFv(UCHT-1)-PE28 FH Key			
		Location/Qualifiers			
		(1)..(1803).			
		FT CDS Location/Qualifiers			
FEATURES	source	1..1803			
		/organism="synthetic construct"			
		/mol_type="genomic DNA"			
		/db_xref="taxon:32630"			
ORIGIN					
	Query Match	67.4%; Score 499.4; DB 6; Length 1803;			
	Best Local Similarity	81.4%; Pred. No. 2.2e-136;			
	Matches 610; Conservative	0; Mismatches 121; Indels 18; Gaps 2;			
QY	1	GATATTGAGATGCACAGACTACATCCCTCGCTGTTCGCCCTCTCGGGACAGAGTCACC 60			
Db	7	GACATCCAGATGACCCAGACCACTCCCTCGCTGTTCGCCCTCTCGGGACAGAGTCACC 66			
QY	61	GTCAATTGCAGGCGAAGTCAGGACATTTAGGAATTATTAACTGGTGATCAGCAGAACA 120			
Db	67	ATCAGTTGACGGCAAGTCAGGACATTTAGGAATTATTAACTGGTGATCAGCAGAACA 126			
QY	121	GACGGAACTGTTAAAATTCCTGATCTACTCACATCAAGATTACTGCCAGAGTCCCATCA 180			
Db	127	GATGGAACTGTTAAAATTCCTGATCTACTCACATCAAGATTACTCAGGAGTCCCATCA 186			
QY	181	AGCTTCAGTCGGCAGTGGGTCTGGAACAGATTAATCCCTCACCATAAACACCTGGAGCAG 240			
Db	187	AAGTTCACTGGCAGTGGGTCTGGAACAGATTAATCTCTCACCATTAGCAACCTGGAGCAA 246			
QY	241	GAAGATATTGCGCACTTACTTTTTGCCAAACGGGCAATACGCTCCGTGGACGTTTCGGTGA 300			
Db	247	GAGGATATTGCCACTTACTTTTTGCCAAACGGGTAATACGCTCCGTGGACGTTTCGGTGA 306			
QY	301	GGACCCAAGCTGGAATAAAACGTGGTGGAGTGGTGTCTGATGGTGGTGGTTCGCGGCG 360			
Db	307	GGACCCAAGCTGGAATCAAACGGGCTGGAGCGGTAGTGGCGGTGGATCGGGTGGAGGC 366			
QY	361	GGCGGCTCCGGTGGTGGATCCGAGGTCCTCAACTGCACAGCTGCAGCTCCTGAGCTGGTG 420			
Db	367	AGC-----GGTGGCGGATCTGAGGTCAGCTCCAGCAGTCTGGA-CCTGAGCTGGTG 417			
QY	421	AAGCTGGGCGCTCAGTGAAGATTTCCTGGAAGAATTTGGCTACGGATTCAAATGCTCT 480			
Db	418	AAGCCTGGAGCTTCAATTGAAGATATCCTGCAAGGCTTCTGTTTACTCATTTCACTGGCTAC 477			
QY	481	TGATGAATCCTGGGTGAAGCAGAGCGCTGGACAGGCTCTTCAGTGGATTTGACCGATTTAT 540			
Db	478	ACCATGNACTGGGTGAAGCAGAGTCATGGAAGAAAGAACTTGAGTGGATGGACCTTTAAT 537			
QY	541	CTTGGAGATGGAGATTTCAACTCAATGGGAAATTCGAGGGCAAGGCCCATCTGCAGCA 600			
Db	538	CTTTACAAGAGTGTGTAGTACCTCAACCAAGATTCAGGCAAGGCCCATTTAACTGTA 597			
QY	601	GACAAATTCCTCCAGCAGCAGCTCATGSCACTCAGCAGCCTGACCTCTGTGGACTCTGG 660			
Db	598	GACAAATTCAGCAGCAGCTCATGGAACCTCTCAGTCTGACATCTGAGGACTCTGCA 657			
QY	661	GTCTATTCTGTGCAAGATCGGGTTGCTACGTTATGCT-----ATGGACTACTGG 711			

Db	658	GTCTATTACTGTGCAAGATCGGGGTACTACGGGTGATAGTGAAGTCTGATCTCGATGCTCGG	717
Qy	712	GGTCACAGGACCTCAGTCAACCGTCTCCTC	740
Db	718	GGCGAGGGACACAGGTCAACCGTCTCCTC	746
RESULT 2			
AX306601			
LOCUS	AX306601	1940 bp	DNA
DEFINITION	Sequence 41 from Patent WO0187982.		linear
ACCESSION	AX306601		
VERSION	AX306601.1	GI:17645800	
KEYWORDS	synthetic construct		
SOURCE	synthetic construct		
ORGANISM	artificial sequences.		
REFERENCE	1		
AUTHORS	Neville,D.M., Thompson,J.T., Hu,H., Woo,J.H., Ma,S., Hexham,J.M.		
TITLE	and Digan,M.E.		
JOURNAL	Immunotoxin fusion proteins and means for expression thereof		
FEATURES	Patent: WO 0187982-A 41 22-NOV-2001;		
	THE DEPARTMENT OF HEALTH & HUMAN SERVICES (US)		
	Location/Qualifiers		
source	1..1940		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="Note = synthetic construct"		
ORIGIN			
Query Match	67.4%;	Score 499.4;	DB 6; Length 1940;
Best Local Similarity	81.4%;	Pred. No. 2.2e-136;	
Matches 610;	Conservative 0;	Mismatches 121;	Indels 18; Gaps 2;
Qy	1	GAATTACAGATGACAGACTACATCCTCCCTGTCTGCTCTCTCTGGGAGACAGAGTCACC	60
Db	1188	GACATCCAGATGACCCAGACACCCTCCTCCCTGTCTGCTCTCTCTGGGAGACAGAGTCACC	1247
Qy	61	GTCAAGTTCAGGCAAGTCAGGACATAGGAATTATTAACTGGTATCAGCAGAAACCA	120
Db	1248	ATCAGTTGACGGCAAGTCAGGACATTTAAATTTAACTGGTATCAACAGAAACCA	1307
Qy	121	GACGGAACCTGTTAAATCCTGATCTACTACACATCAAGATTACTGCCAGAGTCCCATCA	180
Db	1308	GATGGAACCTGTTAACTCCTGATCTACTACACATCAAGATTACTCAGAGTCCCATCA	1367
Qy	181	AGTTTCAGTGGCAGTGGGTCTGGAAACAGATTATTCCTCACCATTAAACCTGGAGCAG	240
Db	1368	AACTTCAGTGGCAGTGGGTCTGGAAACAGATTATTCCTCACCATTAAACCTGGAGCAA	1427
Qy	241	GAAGATTATGGCACTTACTTTTCCCAACAGGCGCAATACGCTCCGTGGAGCTTCCGGTGA	300
Db	1428	GAGGATTATGGCACTTACTTTTCCCAACAGGCTTAACGCTCCGTGGAGCTTCCGGTGA	1487
Qy	301	GGCACCAGCTGAAATAAAACGTGTGGAGTGGTTCATGGTGGTGGTTCCTGGCGGC	360
Db	1488	GGCACCAGCTGAAATCAAAACGGGTGGAGGCGGTAGTGGCGGTGGATCAGGTGGAGGC	1547
Qy	361	GGCGGCTCCGGTGGTGGTGGATCCGAGGTCCTCACTGCAACAGTCTGCACTGAGCTGGTG	420
Db	1548	AGC-----GTTGGCGGATCTGAGGTGCAGCTCCAGCAGTCTGCACTGAGCTGGTG	1598
Qy	421	AAGCCTGGGCGCTCAGTGAAGATTTCCTGCAAAAGATTCTGGCTACGCAATTCAATAGCTCT	480
Db	1599	AAGCCTGGAGCTTCAATGAAGATATCCTGCAAGGCTTCTGGTTACTCACTTCTGCTAC	1658
Qy	481	TGGATCACTGGGTGAAGCAGGCGCTGACAGGCTTCTGAGTGGATTGACCGGATTAT	540
Db	1659	ACCATGAATGGGTGAAGCAGAGTCAATGAAAGAACTTGGAGTGGAGTGAATTAAT	1718
Qy	541	CTTGGAGATGGGATTTCAACTACAATGGGAAATTCGAGGGCAAGGCCATCTGACAGCA	600

Db	1719	CCTTACAAGGTTTAGTACCTACACCAAGAGTTTCAGGACAAGGCCAATTAATCTGTA	1778
Qy	601	GACAAATCCTCCAGCACAGCCTACATGCAGCTCAGCAGCCTGACCTCTGTGGACTCTGG	660
Db	1779	GACAAGTCATCCAGCACAGCCTACATGGAACCTCCTCAGTCTGACATCTGAGGACTCTGCA	1838
Qy	661	GTCTATTCTCTGCACAGATCGGGTGTCTACGTTATCTCT-----ATGGACTACTGG	711
Db	1839	GTCTATTACTGTGCAAGATCGGGGTACTACGGGTAGTAGTACTGGTACTTCGATGTCTGG	1898
Qy	712	GGTCAAGGAACCTCAGTCACCGTCTCCCTC	740
Db	1899	GGCGCAGGACCAAGGTACACCGTCTCCTC	1927
RESULT 3			
AX306600		1943 bp	DNA linear PAT 11-DEC-2001
LOCUS	AX306600	Sequence 40 from Patent WO0187982.	
DEFINITION	AX306600		
ACCESSION	AX306600.1	GI:17645799	
VERSION			
KEYWORDS		synthetic construct	
SOURCE		synthetic construct	
ORGANISM		artificial sequences.	
REFERENCE 1			
AUTHORS	Neville,D.M., Thompson,J.T., Hu,H., Woo,J.H., Ma,S., Hexham,J.M.		
TITLE	and bigan,M.Z.		
JOURNAL	Immunotoxin fusion proteins and means for expression thereof		
FEATURES	Patent: WO 0187982-A 40 22-NOV-2001;		
source	THE DEPARTMENT OF HEALTH & HUMAN SERVICES (US)		
	Location/Qualifiers		
	1..1943		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="Note = synthetic construct"		
ORIGIN			
Query Match	67.4%;	Score 499.4;	DB 6; Length 1943;
Best Local Similarity	81.4%;	Pred. No. 2.2e-136;	
Matches 60;	Conservative 0;	Mismatches 121;	Indels 18; Gaps 2;
Qy	1	GATATTCAGATGACACAGACTACATCCTCCTGCTGCTGCTCTCTGGAGACAGAGTCACC	60
Db	1191	GACATCCAGATGACCCAGACACCTCCTCCTGCTGCTCTCTGGAGACAGAGTCACC	1250
Qy	61	GTCAAGTTCAGGGCAAGTCAGGACATTAGGAATTTAATTAAGTGGTATCAGCAGAAACCA	120
Db	1251	ATCAGTTTCAGGGCAAGTCAGGACATTAGAAATTTATTAACTGGTATCAACAGAAACCA	1310
Qy	121	GACCGAATCTGTTAAATTCCTGATCTACTACACATCAAGATTACTTCGCCAGGAGTCCCATCA	180
Db	1311	GATGGAACTGTTAAACTCCTGATCTACTACACATCAAGATTACTCAGGAGTCCCATCA	1370
Qy	181	AGGTTTCAGTGGCAGTGGGTCTGGACAGATTATTCCTCCATCACCNTAACACCTGGAGCAG	240
Db	1371	AAAGTTCAGTGGCAGTGGGTCTGGAAACAGATTATTTCTTCACCATTAGCAACCTGGAGCAA	1430
Qy	241	GAAGATATTGGCACTTACTTTTGGCAACAGGGCAATACGCCCTCCGTGGACGTTCCGTTGGA	300
Db	1431	GAGGATATTGCCACTTACTTTTGGCCACAGGGTAAACGCTTCGTGGACGTTCCGTTGGA	1490
Qy	301	GGCACCAAGCTGGAAATAAAGCTGGTGGAGGTGGTTCTGATAGTGTGTGTTTCGGCGGC	360
Db	1491	GGCACCAAGCTGGAAATAAAGCTGGTGGAGGTGCACTGCAACAGCTCTGGACCTGAGCTGTG	420
Qy	361	GGCGGCTCCGTTGGTGGTGGATCCGAGGTCCAACTGCAACAGCTCTGGACCTGAGCTGTG	420
Db	1551	AGC-----GGTGGCGGATCTGAGTGCAGCTCCAGCAGCTCTGGACCTGAGCTGTG	1601
Qy	421	AAGCTGGGGCTCAGTGAAGATTCTCTGCAAGATTCTGGCTACGCAATCAATAGTCT	480

Db	1602	AAGCTGGAGCTTCATGAGATATCCTGCAAGCTTCTGGTACTTCATTCACCTGGCTAC	1661
Qy	481	TGGATGAATCGGTGAAGCAGAGCGCTGGACAGGGTCTTGAGTGGATGAGCGGATTTAT	540
Db	1662	ACCATGAATCGGTGAAGCAGAGCTCATGGAAGAACTTGGTGGATGGGACTTATTAAT	1721
Qy	541	CCTGGAGATGGAGTTCCTAACTACATGGGAAATTCGAGGCGAAGCCATACTGACAGCA	600
Db	1722	CCTTACAAGGTGTTAGTACCTACACCGAAGTTCAGAGCAAGGCCATTAACGTGA	1781
Qy	601	GACAAATCCTCCAGCAGCAGCCTACATGCACTCAGCAGCCTGAGCCTCTGTGGACTCTGG	660
Db	1782	GACAAGTCATCCAGCAGCAGCCTACATGGAACCTCCTAGTCTGACATCTGAGGACTCTGCA	1841
Qy	661	GTCATTTCTGTGCAAGATCGGGTGTCTACGTTATGCT-----ATGGACTACTGG	711
Db	1842	GTCATTTACTGTGCAAGATCGGGGTACTACGGGTAGTAGTACTGGTACTTCGATGTCGG	1901
Qy	712	GGTCAAGGAACCTCAGTCACCGTCTCCCTC	740
Db	1902	GGCGCAGGACACCGTCAACCGTCTCCCTC	1930
RESULT 4			
AX306590		2952 bp	DNA
LOCUS	AX306590		linear
DEFINITION	Sequence 30 from Patent WO0187982.		PAT 11-DEC-2001
ACCESSION	AX306590		
VERSION	AX306590.1	GI:17645796	
KEYWORDS	synthetic construct		
SOURCE	synthetic construct		
ORGANISM	artificial sequences.		
REFERENCE	1.		
AUTHORS	Neville,D.M., Thompson,J.T., Hu,H., Woo,J.H., Ma,S., Hexham,J.M. and Bigan,M.E.		
TITLE	Immunotoxin fusion proteins and means for expression thereof		
JOURNAL	Patent: WO 0187982-A 30 22-NOV-2001;		
FEATURES	THE DEPARTMENT OF HEALTH & HUMAN SERVICES (US)		
source	Location/Qualifiers		
	1..2952		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="Note = synthetic construct"		
ORIGIN			
Query Match	65.2%;	Score 482.8;	DB 6; Length 2952;
Best Local Similarity	80.2%;	Pred. No. 1.9e-131;	
Matches 587;	Conservative 0;	Mismatches 127;	Indels 18; Gaps 1
Qy	1	GATATTCAGATGACACAGACTACATCCTCCCTGTCTGCTCCTCTCTGGGAGACAGATCACC	60
Db	1435	GACATCCAGATGACCCAGACACCTCCTCCCTGTCTGCTCCTCCCTGGGCGACAGAGTCACC	1499
Qy	61	GTCAATTGCAGGCGAGTCAGGACATTAGGAATTTAACTGGTATCAGCAGAAACCA	120
Db	1495	ATCAGTTGCAGGCGAGTCAGGACATTAGAAATTTAACTGGTATCAACAGAAACCA	1555
Qy	121	GACCGAATCTTAAATTCCTGATCTACTACATCAAGATTACTGCCAGGAGTCCCATCA	180
Db	1555	GATGGAACTGTTAACTCCTGATCTACTACATCAAGATTACTCAGGAGTCCCATCA	161
Qy	181	AGGTTTCAGTGCAGTGGGTCTGGAAACAGATTATTCCTCACCATTAACTGAGCAG	240
Db	1615	AAGTTTCAGTGCAGTGGGTCTGGAAACAGATTATTCCTCACCATTAGCAACCTGGAGCA	167
Qy	241	GAAGATATTGGCACTTACTTTTGGCAACAGGGCAATACGCTTCGTTGACGTTTCGTTGA	300
Db	1675	GAGGATATTGCCATTACTTTTGGCAACAGGGTATACGCTTCGTTGACGTTTCGTTGA	173
Qy	301	GGCAACCAAGCTGGAATAAAACGTTGGTGGAGGTGGTCTGTGATGGTGGTGGTTCGGCGGC	360

Db 1735 GGCACCAAGCTGGAGATAAAAGGAGGCGGAGGCAG-----CGGAGGC 1776
Qy 361 GCGGCTCCGGTGGTGGATCCGAGTCCAACTGCAACAGTCTGGACCTGAGCTGGTG 420
Db 1777 GGTGGCTCGGAGGCGGAGGCTCGAGTGCAGCTCCAGCAGTCTGGACCTGAGCTGGTG 1836
Qy 421 AAGCCTGGGCTCAGTGAAGATTCTGCAAGATTCTGGACGCAATCAATAGCTCT 480
Db 1837 AAGCCTGGAGCTTCAATGAAGATATCTGCAAGGCTTCTGTTACTCAATCACTGGCTAC 1896
Qy 481 TGGATGAAGTGGGGAAGCAGAGGCTGGACAGGCTCTGAGTGGATTGGACGGATTAT 540
Db 1897 ACCATGAAGTGGGGAAGCAGAGTCTGGAAGAACCTTGGTGGATGGAGCTTATTAAT 1956
Qy 541 CTTGAGATGGAGATTCTAACTACAATGGGAAATTCGAGGCAAGGCCATCTACAGCA 600
Db 1957 CTTTACAAAGGTGTAGTACTCAACCCAGAGTTCAAGGACAGGCCACATTAATCTGA 2016
Qy 601 GACAAATCTCCAGCAGCCTACATGCACTGAGCTGAGCAGCTGAGCTCTGGACTCTGG 660
Db 2017 GACAAATCTCCAGCAGCCTACATGCACTGAGCTGAGCAGCTGAGCTCTGGACTCTGG 2076
Qy 661 GTCTATTCTGTGCAAGATCGGGTGTGCTAGTATGCTATGGACTACTGGGCTCAAGGA 720
Db 2077 GTCTATTACTGTGCAAGATCGGGTGTGCTAGTATGCTATGGACTACTGGGCTCAAGGA 720
Qy 721 ACTCAGTCAACC 732
Db 2137 GCGCAGGACCC 2148

RESULT 5
AR116105 916 bp DNA linear PAT 16-MAY-2001
LOCUS AR116105
DEFINITION Sequence 29 from patent US 6132992.
ACCESSION AR116105
VERSION AR116105.1 GI:14096427
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 916)
AUTHORS Ledbetter,J.A., Gilliland,L.K., Hayden,M.S., Linsley,P.S.,
Bajorath,J. and Fell,H.Perry.
TITLE Expression vectors encoding bispecific fusion proteins and methods
of producing biologically active bispecific fusion proteins in a
mammalian cell
JOURNAL Patent: US 6132992-A 29 17-OCT-2000;
FEATURES Location/Qualifiers
source 1..916
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 64.7%; Score 479.2; DB 6; Length 916;
Best Local Similarity 79.4%; Pred. No. 2e-130;
Matches 581; Conservative 0; Mismatches 148; Indels 3; Gaps 1;
Qy 1 GATATTAGATGACACAGACTACATCTCCCTGTCTGCTCTCTGGGAGACAGAGTCAACC 60
Db 76 GACATCCAGATGACACAGACTACATCTCCCTGTCTGCTCTCTGGGAGACAGAGTCAACC 135
Qy 61 GTCAGTTGAGGCGAGTACAGATAGGATTTAACTGATATGATGATGATGATGATGATGAT 120
Db 136 ATCAGTTGAGGCGAGTACAGATAGGATTTAACTGATATGATGATGATGATGATGATGAT 195
Qy 121 GACGGAATGTTAAATTCCTGATCTACTACATCAAGATTACTGCCAGGAGTCCCATCA 180
Db 196 GATGGAATGTTAAATTCCTGATCTACTACATCAAGATTACTGCCAGGAGTCCCATCA 255
Qy 181 AGGTTGAGTGGAGTGGTGTGGAACAGATTATTCCTCACCATTAACTGATGATGATGATGAT 240

Db 256 AGGTTCACTGGCAGTGGGTCTGGAAACAGATTATTTCTCTACCAATTGCCAACCTGCAACCA 315
Qy 241 GAAGATATTGGCACTTACTTTTGGCAACAGGCAATACGCTCCGCTGGACCTTCGGTGGGA 300
Db 316 GAAGATATTGGCACTTACTTTTGGCAACAGGCAATACGCTTCGCTGGACCTTCGGTGGGA 375
Qy 301 GGCACCAAGCTGGAAATAAACAGTGGTGGAGTGGTCTGATGTTGGTGGTCTTGGGGGC 360
Db 376 GGCACCAAGCTGGTAAACCAACAGGAGCTCGGTGGCGGTGGCTCGGGCGGTGGTGGTGG 435
Qy 361 GCGGCTCCGGTGGTGGTGGATCCGAGGTCCAACTGCAACAGTCTGGACCTGAGTGGTGG 420
Db 436 GGTGG---CGGCGGATCTATCGATGAGGTCCAGCTGCAACAGTCTGGACCTGAGTGGTGG 492
Qy 421 AAGCCTGGGCTCAGTGAAGATTCTTGGAAAAGATTCTGGCTAGCAGATTCAATAGCTCT 480
Db 493 AAGCCTGGAGCTTCAATGACAAATGCTTGCAGGCTCTGTTACTCAATCACTGGCTAC 552
Qy 481 TGGATGAAGTGGGTGGAAGCAGAGGCTCGACAGGCTCTGAGTGGATTGGACGGATTAT 540
Db 553 ATCTGTGAAGTGGTGAAGCAGAGCTTGAAGAACCTTGGTGGATTGGACTTATTAAT 612
Qy 541 CTTGAGATGGAGATTCTAACTACAATGGGAAATTCGAGGCAAGGCCATCTACAGCA 600
Db 613 CCATACAAAGGTCTTACTACCTTACCAACAGAAATTCAGGCGCAAGGCCACATTAATCTGA 672
Qy 601 GACAAATCTCCAGCAGCCTACATGCACTGAGCTGAGCAGCTGAGCTCTGGACTCTGG 660
Db 673 GACAAATCTCCAGCAGCCTACATGAGCTCTGAGTCTGACATCTGAGACATCTGAGAC 732
Qy 661 GTCTATTCTGTGCAAGATCGGGTGTGCTAGTATGCTATGGACTACTGGGCTCAAGGA 720
Db 733 GTCTATTACTGTGCAAGATCTGGTACTATGGTACTCGGACTGCTGATCTTTCGATGCTGG 792
Qy 721 ACTCAGTCAACC 732
Db 793 GCGCAGGACCC 804

RESULT 6
145604 916 bp DNA linear PAT 07-OCT-1997
LOCUS 145604
DEFINITION Sequence 29 from patent US 5637481.
ACCESSION 145604
VERSION 145604.1 GI:2469706
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 916)
AUTHORS Ledbetter,J.A., Gilliland,L.K., Hayden,M.S., Linsley,P.S.,
Bajorath,J. and Fell,H.Perry.
TITLE Expression vectors encoding bispecific fusion proteins and methods
of producing biologically active bispecific fusion proteins in a
mammalian cell
JOURNAL Patent: US 5637481-A 29 10-JUN-1997;
FEATURES Location/Qualifiers
source 1..916
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 64.7%; Score 479.2; DB 6; Length 916;
Best Local Similarity 79.4%; Pred. No. 2e-130;
Matches 581; Conservative 0; Mismatches 148; Indels 3; Gaps 1;
Qy 1 GATATTAGATGACACAGACTACATCTCCCTGTCTGCTCTCTGGGAGACAGAGTCAACC 60
Db 76 GACATCCAGATGACACAGACTACATCTCCCTGTCTGCTCTCTGGGAGACAGAGTCAACC 135
Qy 61 GTCAGTTGAGGCGAGTACAGATAGGATTTAACTGATATGATGATGATGATGATGATGAT 120
Db 136 ATCAGTTGAGGCGAGTACAGATAGGATTTAACTGATATGATGATGATGATGATGATGAT 195

QY 121 GAGGAACTGTTAAATTCCTGATCTACTACACATCAAGATTACTGCCAGGAGTCCCATCA 180
Db 196 GATGGAACTGTTAAATTCCTGATCTACTACACATCAAGATTACTGCCAGGAGTCCCATCA 255
QY 181 AGGTTAGTGGCAGTGGTCTGGACAGATATTCCTCAGCATTAACACCTGGAGCAG 240
Db 256 AGGTTAGTGGCAGTGGTCTGGACAGATATTCCTCAGCATTAACACCTGGAGCAG 315
QY 241 GAAGATATTGGCACTTACTTTTGGCAACAGGCAATACGCTTCGGTGGG 300
Db 316 GAAGATATTGGCACTTACTTTTGGCAACAGGCAATACGCTTCGGTGGG 375
QY 301 GGCAACAGCTGGGTAATAAAGCTGGTGGAGTGGTCTGATGGTGGTCTGGCGGC 360
Db 376 GGCAACAGCTGGGTAATAAAGCTGGTGGAGTGGTCTGATGGTGGTCTGGCGGC 435
QY 361 GGCGGCTCGGTTGGTGGTGGAGTGGTGGAGTGGTGGAGTGGTGGAGTGGTGGAG 420
Db 436 GGTGG---CGGCGGATCTATCGATGAGTGGTGGAGTGGTGGAGTGGTGGAGTGG 492
QY 421 AAGCTTGGGCTCAGTGAAGATTTCTGCAAGATTTCTGGTACGCAATTAAGTCT 480
Db 493 AAGCTTGGGCTCAGTGAAGATTTCTGCAAGATTTCTGGTACGCAATTAAGTCT 552
QY 481 TGGATGAAGCTGGTGGTGGAGTGGTGGAGTGGTGGAGTGGTGGAGTGGTGGAG 540
Db 553 ATCGTGAAGCTGGTGGTGGAGTGGTGGAGTGGTGGAGTGGTGGAGTGGTGGAG 612
QY 541 CCTGGAGATGAGATTTCTAACTAAGTGGGAAATTCGAGGGCAAGGCCATTAAGTGA 600
Db 613 CCATACAAAGTCTTACTACCAAGTGGGAAATTCGAGGGCAAGGCCATTAAGTGA 672
QY 601 GACAAATCCCGACAGCAGCTACATGAGCTGAGCAGCTGAGCTGAGCTGAGCTGCG 660
Db 673 GACAAATCCCGACAGCAGCTACATGAGCTGAGCAGCTGAGCTGAGCTGAGCTGCG 732
QY 661 GTCTATTCTGTGCAAGATCGGGTGTCTAGTATGCTATGAGTACTGGGTCAGGA 720
Db 733 GTCTATTCTGTGCAAGATCGGGTGTCTAGTATGCTATGAGTACTGGGTCAGGA 792
QY 721 ACCTCAGTCACC 732
Db 793 GCGCAGGGGACC 804

RESULT 7
AR256099 LOCUS 916 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 29 from patent US 6482919.
ACCESSION AR256099
VERSION AR256099.1 GI:27305471
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 916)
AUTHORS Ledbetter,J.A., Hayden,M.S., Linsley,P.S., Bajorath,J., Fell,H.P.
and Gilliland,L.K.
TITLE Expression vectors encoding bispecific fusion proteins and methods of producing biologically active bispecific fusion proteins in a mammalian cell
JOURNAL Patent: US 6482919-A 29 19-NOV-2002;
FEATURES Location/Qualifiers
source 1..916
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 64.7%; Score 479.2; DB 6; Length 916;
Best Local Similarity 79.4%; Pred.No.2e-130;
Matches 581; Conservative 0; Mismatches 149; Indels 3; Gaps 1;

QY 1 GATATTGAGTGAACAGACTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
Db 76 GACATCCAGATGACACAGACTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 135
QY 61 GTCAGTTCAGGGCAAGTCAAGCAATTTAGGAATTTTAAACTGGTATCAGCAGAAACCA 120
Db 136 ATCAGTTCAGGGCAAGTCAAGCAATTTAGGAATTTTAAACTGGTATCAGCAGAAACCA 195
QY 121 GACGAACTGTTAAATTCCTGATCTACTACATCAAGATTACTGCCAGGAGTCCCATCA 180
Db 196 GATGAACTGTTAAATTCCTGATCTACTACATCAAGATTACTGCCAGGAGTCCCATCA 255
QY 181 AGGTTAGTGGCAGTGGTCTGGACAGATATTCCTCAGCATTAACACCTGGAGCAG 240
Db 256 AGGTTAGTGGCAGTGGTCTGGACAGATATTCCTCAGCATTAACACCTGGAGCAG 315
QY 241 GAAGATATTGGCACTTACTTTTGGCAACAGGCAATACGCTTCGGTGGG 300
Db 316 GAAGATATTGGCACTTACTTTTGGCAACAGGCAATACGCTTCGGTGGG 375
QY 301 GGCAACAGCTGGGTAATAAAGCTGGTGGAGTGGTGGAGTGGTGGAGTGGTGGAG 360
Db 376 GGCAACAGCTGGGTAATAAAGCTGGTGGAGTGGTGGAGTGGTGGAGTGGTGGAG 435
QY 361 GGCGGCTCGGTTGGTGGTGGAGTGGTGGAGTGGTGGAGTGGTGGAGTGGTGGAG 420
Db 436 GGTGG---CGGCGGATCTATCGATGAGTGGTGGAGTGGTGGAGTGGTGGAGTGG 492
QY 421 AAGCTTGGGCTCAGTGAAGATTTCTGCAAGATTTCTGGTACGCAATTAAGTCT 480
Db 493 AAGCTTGGGCTCAGTGAAGATTTCTGCAAGATTTCTGGTACGCAATTAAGTCT 552
QY 481 TGGATGAAGCTGGTGGTGGAGTGGTGGAGTGGTGGAGTGGTGGAGTGGTGGAG 540
Db 553 ATCGTGAAGCTGGTGGTGGAGTGGTGGAGTGGTGGAGTGGTGGAGTGGTGGAG 612
QY 541 CCTGGAGATGAGATTTCTAACTAAGTGGGAAATTCGAGGGCAAGGCCATTAAGTGA 600
Db 613 CCATACAAAGTCTTACTACCAAGTGGGAAATTCGAGGGCAAGGCCATTAAGTGA 672
QY 601 GACAAATCCCGACAGCAGCTACATGAGCTGAGCAGCTGAGCTGAGCTGAGCTGCG 660
Db 673 GACAAATCCCGACAGCAGCTACATGAGCTGAGCAGCTGAGCTGAGCTGAGCTGCG 732
QY 661 GTCTATTCTGTGCAAGATCGGGTGTCTAGTATGCTATGAGTACTGGGTCAGGA 720
Db 733 GTCTATTCTGTGCAAGATCGGGTGTCTAGTATGCTATGAGTACTGGGTCAGGA 792
QY 721 ACCTCAGTCACC 732
Db 793 GCGCAGGGGACC 804

RESULT 8
AR401482 LOCUS 916 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 29 from patent US 6623940.
ACCESSION AR401482
VERSION AR401482.1 GI:40148818
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 916)
AUTHORS Ledbetter,J.A. and Linsley,P.S.
TITLE Expression vectors encoding bispecific fusion proteins and methods of producing biologically active bispecific fusion proteins in a mammalian cell
JOURNAL Patent: US 6623940-A 29 23-SEP-2003;
FEATURES Location/Qualifiers
source 1..916
/organism="unknown"
/mol_type="genomic DNA"

[illegible]

Db 469 AAGCCTCGGGGTCCTGTAACCTCTCTGTGCAGCCTCTGACTCACTTTCACTAGCTAT 528
Qy 481 TGGATGAACCTGGGTGAAGACAGAGCCCTGGACAGGGTCTTGAGTGGATGAGCGGATTTAT 540
Db 529 GCCATGCTTGGGTTCGCCAGACTCCAGAGAGAGGCTGGAGTGGGTGCGATCCATTAGT 588
Qy 541 CCTGGAGATGGAGATTCTAACTACAATGGGAATTCGAGGGCAAGGCCATCTACGACGCA 600
Db 589 AGTGG---TGGTTTCACTACTATCCAGACAGTGTGAAGGCCGATTCACCATCTCCAGA 645
Qy 601 GACAAATCTCCACACAGCCTACATGAGCTCAGCAGCTCAGCCTGAGCTCTGTGGACTCTGGG 660
Db 646 GATAATGCCAGGAACATCTGTATCTGCAATGAGCAGTCTGAGTCTGAGGACACGGCC 705
Qy 661 GTCTATTCTGTGCAAGATCGGGTCTGATGTTATGCTATGAGTACTGAGGCTCAGGA 720
Db 706 ATGTATTACTGTGCAAGAGACAGGTACGGGGTA---CCTCGATGCTCGGGCGCAGGG 762
Qy 721 ACCTCAGTCACCGTCTCCTCG 741
Db 763 ACCAGGTCACCGTTTCCTCG 783

RESULT 12
AX797327
LOCUS AX797327 1956 bp DNA linear PAT 04-OCT-2003
DEFINITION Sequence 4 from Patent WO03051926.
ACCESSION AX797327
VERSION AX797327.1 GI:37517969
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Fey, G.H., Gramatzki, M. and Peipp, M.
TITLE Anti-cd7 immunotoxin as fusion protein
JOURNAL Patent: WO 03051926-A 4 26-JUN-2003;
Friedrich-Alexander-Universitaet Erlangen-Nuernberg (DE)
FEATURES
source
1. 1956
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 56.8%; Score 421; DB 6; Length 1956;
Best Local Similarity 74.9%; Pred. No. 3.7e-113;
Matches 555; Conservative 0; Mismatches 180; Indels 6; Gaps 2;

Qy 1 GATATTTCAGTACACAGACTACATCCCTCCTGCTGCTCTCTGGGAGACAGAGTCACC 60
Db 49 GATATCCAGATGACAGACTACATCCCTCCTGCTGCTCTCTGGGAGACAGAGTCACC 108
Qy 61 GTCAGTTGAGGGCAAGTCAGACATTTAGGAATTTAACTGTTATCAGCAGAAACCA 120
Db 109 ATCAGTTGAGTGCAGTCAGGCAATTTAGCAATTTAACTGTTATCAGCAGAAACCA 168
Qy 121 GACGAGCTGTTAAATCTCTGATCTACTACATCAAGATTACTCCAGGAGTCCATCA 190
Db 169 GATGGACCTGTTAAATCTCTGATCTATTACATCAAGTTTACACTAGGAGTCCATCA 228
Qy 181 AGTTTCAGTGGCAGTGGGTCTGGAACAGATTATTCCTCACCANTPAA CAACCTGGAGAG 240
Db 229 AGTTTCAGTGGCAGTGGGTCTGGAACAGATTATTCCTCACCANTCAGCAACCTGGAACCT 288
Qy 241 GAAGATTATGGCACTTACTTTTGCAACAGGCAATAGCCTCGGTGGAGCTTCGGTGA 300
Db 289 GAAGATTATGGCACTTATTATTGTGACAGATATAGCAAGCTTCGGTACAGCTTCGAGGG 348
Qy 301 GGCACCAAGCTGGAATAAAGCTGGTGGAGTGGTCTTGATGGTGGTCTTCGGCGGC 360
Db 349 GGGACCAAGCTGGAATAAAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 408

Qy 361 GCGGCTCCGGTGGTGGATCCAGTCCAACTGCAACAGTCTTGACCTGAGCTGGTG 420
Db 409 GCGGCTCCGGTGGTGGATCCAGTCCAACTGCTGGAGTCTGGGGAGGCTTAGTG 468
Qy 421 AAGCCTCGGGCTCAGTGAAGATTTCCTGCAAGATTCTGGCTACGCAATTCATAGCTCT 480
Db 469 AAGCCTCGGGGTCCTGAAACTCTCTGTGCGAGCCTCTGGACTCACTTTCACTAGTAT 528
Qy 481 TGGATGAACCTGGGTGAAGACAGAGCCTGGACAGGGTCTTGAGTGGATTTGACGGATTAT 540
Db 529 GCCATGCTTGGGTTCGCCAGACTCCAGAGAGAGGCTGGAGTGGGTGCGATCCATTAGT 588
Qy 541 CTTGAGATGAGAGATTCTAATCTACAAATGGGAAATTCGAGGGCAAGGCCATCTACAGCA 600
Db 589 AGTGG---TGGTTTCACTACTATCCAGACAGTGTGAAGGGCCGATTCACCATCTCCAGA 645
Qy 601 GACAAATCTCCACACAGCCTACATGAGCTCAGCAGCTCAGCAGCTCTGTGGACTCTCGG 660
Db 646 GATAATGCCAGGAACATCTGTATCTGCAATGAGCAGTCTGAGTCTGAGGACACGGCC 705
Qy 661 GTCTATTCTGTGCAAGATCGGGTTCCTACGTTATGCTATGAGTACTGAGGCTCAAGGA 720
Db 706 ATGTATTACTGTGCAAGAGACAGGTACGGGGTA---CCTCGATGCTCGGGCGCAGGG 762
Qy 721 ACCTCAGTCACCGTCTCCTCG 741
Db 763 ACCAGGTCACCGTTTCCTCG 783

RESULT 13
E35160
LOCUS E35160 1457 bp DNA linear PAT 31-JAN-2002
DEFINITION Method for inducing apoptosis.
ACCESSION E35160
VERSION E35160.1 GI:18624419
KEYWORDS JP 2000102389-A/7.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1457)
AUTHORS Hagiwara, H., Aotsuka, Y. and Miyahara, J.
TITLE Method for inducing apoptosis
JOURNAL Patent: JP 2000102389-A 7 11-APR-2000;
YOSHIIDE HAGIWARA, HIDEAKI HAGIWARA
COMMENT OS Artificial Sequence
PN JP 2000102389-A/7
PD 11-APR-2000
PF 29-SEP-1998 JP 1998291441
PR

PI HIDEAKI HAGIWARA, YASUYUKI AOTSUKA, JUNICHI MIYAHARA PC
C12N15/09, A61K31/00, A61K48/00, C07K16/42, C07K19/00, C12N5/10, PC
C12N15/02//
PC (C12N5/10, C12R1:91), C12N5/00, C12N5/00, C12N15/00, C12N5/00, PC
C12R1:91)

CC
FH Key Location/Qualifiers
FT CDS

FEATURES
source
1. 1457
Location/Qualifiers
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 56.5%; Score 418.6; DB 6; Length 1457;
Best Local Similarity 73.6%; Pred. No. 1.8e-112;
Matches 554; Conservative 0; Mismatches 184; Indels 15; Gaps 1;
Qy 1 GATATTTCAGTACAGACTACATCCCTCCTGCTGCTCTCTGGGAGACAGAGTCACC 60
Db 67 GACATCCAGATGAACAGTCTCCATCAGTCTCTGTCATCCCTTGGAGACACATATCC 126

Qy	61	GTCA	GTGTCAGGGCAAGTCACGACATTAAGGAATTA	TTAACTCGGTATCAGCAGAAACCA	120
Db	127	ATCA	CTTGCCATGCCAGTCAGAA	CATTGATGTTTGGTTAAAGCTGGTACACGACGAAACCA	186
Qy	121	GACG	GAACCTGTTTAAATTCCTGATCTACTACACATCAAGATTACTGC	CAGAGTCCCATCA	180
Db	187	GGAA	TGTTCCTPAACTATTGATCTATAAGGCTTCCAACTTG	CACACAGGCGTCCCATCA	246
Qy	181	AGGT	TCAGTGGCAGTGGGTCTGGAAACAGATTATTCCTCCACCATTAACAACTCTGGAGCAG	240	
Db	247	AGGT	TTAGTGGCAGTGGATCTGGAACAGGTTTCAATTAACCATCAGCAGCCTGCAGCCT	306	
Qy	241	GAAGA	TATTGGCACATTACTTTTGGCCAACAGGGCAATACGCTCGTGGAGAGTTTCGGTGA	300	
Db	307	GAAGA	CATTGCCACATTACTATCTGTCCACAGAGGTCAAGTTATCTCGAAGTTTCGGTGA	366	
Qy	301	GGCA	CAAGCTCGGAATAAAACGTGGTGGAGGTGGTTT	-----TGATGT	345
Db	367	GGCA	CAAGCTCGGAATAAAACGGGCTGATGCTGCACCAACTGTATCCCTCGAGGGTGA	426	
Qy	346	GGTGG	TTCTGGCGGGCGGCTCCGGTGGTGGATCCGAGGTCACACTCGAAACAGTCT	405	
Db	427	GGCG	GTTCAGGGCGGAGTGGCTCTCGCCGGTGGCGGATCAGAGGTTCAGCTTCAGACAGTCT	486	
Qy	406	GGAC	CTGAGCTGGTGAAGCCTGGGGCCTCAGTGAAGATTTCCTGCCAAAGATTCTGGCTAC	465	
Db	487	GGCA	CTGTCTGCAAGGCCTGGGGCTTCAGTGAAGATGTCCTCGAAGGCTTCGGGCTAC	546	
Qy	466	GCA	TTCAATAGCTCTTGGATGAATCTGGGTGAAGCAGAGCCCTGCACAGGGCTTGTAGTGG	525	
Db	547	ACCT	TTAAACAGCTACTCTGAGATGACCTGGGGTAAACAGAGCCCTGGACAGGGTCTGGAAATGG	606	
Qy	526	ATTG	GACGATTATTCCTCGAGATGGAGATTCTACTACAATGGGAAATTCGAGGGCAAG	585	
Db	607	ATTG	GGCGGATTATTCCTGGAAATAGTGATATTAGCTACAGCCAGAACTTTAAGGACAGG	666	
Qy	586	GCCAT	CTGACGACGACAAATCTCTCAGCACAGCCTACATGACGCTCAGCAGCCTGACC	645	
Db	667	GCCAA	ACTGACTCCCGTCACTCCACCAGCACTGGCTCATATGGAACTCAGAAAGCCTGACA	726	
Qy	646	TCGT	GTGACTCTCGGGTCTATTTCGTGCAAGATCGGGTGTACGTTATGCTTATGGAC	705	
Db	727	AATG	AGACTCTCGGGTCTATTCTGTGTACAAAGAGGAATATGATACGACACCTGGAC	786	
Qy	706	TACT	GGGGTCAAGGAACCTCAGTCAACCGTCTCC	738	
Db	787	TACT	GGGTCAAGGAACCTCAGTCAACCGATCC	819	

RESULT 14	AX398258	LOCUS	AX398258	1545 bp	DNA	linear	PAT 27-MAY-2002
		DEFINITION	Sequence 17 from Patent WO0220615.				

ORGANISM	Mus sp.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus	
REFERENCE	
1. Mack M., Schloendorff D. and Spring M.	

```

FEATURES             source
Micromet AG (DE)    Location/Qualifiers
1..1545
    /organism="Mus sp."
    /mol_type="unassigned DNA"
    /db_xref="taxon:10095"
<58..1545
    /note="unnamed protein product"

```

Search completed: April 26, 2004, 16:55:18
Job time : 3248 secs